Simultaneous Nonlinear Estimation

John J. Beauchamp

February 1, 1967

~~Technical Report Number~~-11

NASA Grant Number NGR-10-004-029

Department of Statistics
Florida State University
Tallahassee, Florida

ACKNOWLEDGEMENTS

I am most deeply indebted to Dr. Richard G. Cornell, my major professor, who introduced me to the general problem of nonlinear estimation, and patiently guided me during the research on the particular problem of simultaneous nonlinear estimation contained in this dissertation. Besides the helpful suggestions from Dr. Cornell, I also wish to acknowledge the assistance from the other members of my Supervisory Committee: Drs. H. C. Griffith, S. K. Katti, D. A. Meeter, and W. A. Thompson, Jr.

I wish to acknowledge the financial assistance that I received while I was enrolled at Florida State University from a Public Health Service Fellowship (5-F1-GM-15,632-03), during the period 1962-65, a training grant from the National Institute of General Medical Sciences (5T1 GM-913), and a NASA grant (NGR-10-004-029), during the period 1965-66. In addition, I want to recognize a grant from the National Science Foundation to the Florida State University Computing Center, which assisted with the computations of this research.

I wish to acknowledge the faculty of the Department of Statistics at Florida State University for their courses and assistance during my enrollment at the University, and also the many graduate students in the Department of Statistics for their helpful informal discussions related to this research. I want to thank the secretarial staff of the Department of Statistics for doing a very good job of typing a difficult manuscript. As always, the responsibility for any errors in this dissertation rests with me, the author.

John J. Beauchamp

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

In the fields of biology, chemistry, and physics there are numerous examples of experimental situations which yield data that may be reasonably described by the following set of regression equations:

$$Y_{ij} = \alpha_{io} + \sum_{k=1}^{m} \alpha_{ik} e^{-\lambda_k x_j} + \varepsilon_{ij} \tag{1.1}$$

for $i = 1,2,\ldots,n$ and $j = 1,2,\ldots,N$. In this expression Y_{ij} and ε_{ij} represent random variables associated with the j^{th} observation on the i^{th} equation; x_j represents an independent variable, e.g., time; and the α_{ik}'s and λ_k's are constant parameters that are inherent in the physical or experimental situation. One of the prime examples of such an experimental situation which will be referred to a great number of times during this research involves the use of radioactive tracers in order to study certain biological processes.

For single equation regression models, there have been many estimation procedures developed and in Chapter 2 we will discuss some of the single equation nonlinear estimation procedures that are pertinent to this research as well as other related literature. In Chapter 3 we will present a discussion of the various types of regression models to be used in this research and we will show how the regression equations used to describe a mammillary or catenary compartment model are members of the class of regression models given by (1.1). Also in this chapter some of the distributional assumptions concerning the random variables will be stated. In Chapter 4 a generalized least squares estimation procedure will be presented and evaluated. This generalized procedure may be applied not only to the class of regression models given by (1.1), but, under certain regularity conditions to be given then, also to the more general class of regression models described by:

$$Y_{ij} = f_i(X_j;\theta) + \varepsilon_{ij} \tag{1.2}$$

where Y_{ij} and ε_{ij} are the same as defined in (1.1); f_i represents the regression function for the i^{th} equation; X_j represents a vector of independent variables; and θ represents a vector of constant parameters to be estimated. In Chapter 5 a generalization of the partial total estimation procedure as discussed by Cornell [1962] will be presented and evaluated. This estimation procedure will be applied to the class of regression functions in (1.1) when $n = m$. In Chapter 6 a generalization of the Spearman estimation procedure as discussed by Johnson and Brown [1961] will be presented and evaluated for a particular class of regression functions in (1.1) when $n = m$.

From an investigation of equation (1.1) we note that the set of exponential parameters $\lambda_1,\lambda_2,\ldots,\lambda_m$ is present in each one of the regression equations. These exponential parameters are usually of primary interest in experimental situations and therefore our main concern has been with the estimation of the parameters that are common to some of the regression equations, even though our general concern has been with the estimation of all of the constant parameters in our regression model. We also have dealt exclusively with estimation schemes that make use simultaneously of all of the observations available on all of the equations being studied. In Chapters 4, 5, and 6 the range of the subscript j of equations (1.1) and (1.2) will change as we go from one estimation scheme to another, since the requirements on the total number of observations and the spacing of the observations change from one procedure to another. However, we will attempt to make these changes clear as the different procedures are presented.

In Chapter 7 numerical examples of the various estimation procedures will be given and a comparison of the various procedures will be given. This chapter will also contain suggestions that will help the experimenter design his experiment in order to make the best use of these estimation procedures.

II. REVIEW OF LITERATURE

2.1 Introduction

In this chapter we want to review papers that are concerned with presently used nonlinear estimation procedures before our generalized estimation schemes are presented. In Section 2.2 our discussion will be devoted to those papers that consider the problem of nonlinear estimation for a single nonlinear regression equation. Also included will be a discussion of least squares nonlinear estimation as well as other nonlinear estimation procedures. In Section 2.3 papers that have considered the problem of simultaneous estimation of parameters in a set of regression functions will be discussed. Then since some of the generalized estimation procedures presented in this research are restricted to particular regression models, Section 2.4 will discuss how the models as given by equation (1.1) arise in tracer experiments. The papers referred to in this chapter are meant in no way to be a complete review of the literature concerned with nonlinear estimation; however, they have been selected because of their pertinence to this research.

2.2 Single equation estimation procedures

In this section we will be interested only in the case when $n = 1$ from equation (1.2), and therefore we will suppress the subscript i.

One of the more heavily used nonlinear estimation procedures is the iterative least squares technique which may be attributed to the work of Gauss (see translation of Gauss' work by Trotter [1957]), who was among the first to use the Newton or Newton-Raphson method for the specific purpose of estimating the parameters in nonlinear regression equations by the least squares method. The basic aim of the least squares estimation techniques is the minimization of

$$\phi(\theta) = \sum_{j=1}^{N} [y_j - f(X_j;\theta)]^2 \tag{2.1}$$

where y_j, f, X_j, and θ have been defined in equation (1.2). The usual Gaussian procedure involves the approximation of the original nonlinear function, f, by a linear model by means of a Taylor series expansion of f through the linear terms about a point of initial estimates of the vector of parameters, θ. The details of this estimation procedure will be presented in Chapter 4 since this procedure is incorporated into the simultaneous nonlinear least squares estimation method developed there.

One of the problems that occasionally arises with the application of the Gaussian iterative technique of estimation is the problem of convergence, and some authors have presented modifications to the Gaussian technique in order to circumvent this problem. Hartley [1961] presents some assumptions for and modifications to the Gauss-Newton estimation method and his modified Gauss-Newton method has the merit of guaranteed convergence under the assumptions to be given in Chapter 4 where the detailed steps of the modified procedure will also be given. Levenberg [1944] extends the standard Gaussian iterative technique so that an improvement of the initial estimate, ${}_o\hat{\theta}$, of the vector of parameters θ could be ensured, i.e. improved in the sense that $\phi({}_1\hat{\theta}) < \phi({}_o\hat{\theta})$ where ${}_1\hat{\theta}$ is the new estimate of the vector θ derived from the initial estimate ${}_o\hat{\theta}$. Levenberg proposes that the following augmented sum of squares be minimized over the range of the elements of θ:

$$\bar{\phi}^* = w\phi^* + a_1\, {}_o\hat{\delta}_1^2 + a_2\, {}_o\hat{\delta}_2^2 + \ldots + a_p\, {}_o\hat{\delta}_p^2 \tag{2.2}$$

where

$$\phi^* = \sum_{j=1}^{N} (y_j - {}_of_j - {}_o\hat{\delta}_1\, {}_of_j^{(1)} - \ldots - {}_o\hat{\delta}_p\, {}_of_j^{(p)})^2, \tag{2.3}$$

${}_o\hat{\delta}_b = (\theta_b - {}_o\hat{\theta}_b)$; ${}_of_j = f(X_j;{}_o\hat{\theta})$; and ${}_of_j^{(b)} = \dfrac{\partial f(X_j;{}_o\hat{\theta})}{\partial\theta_b}$ for $b = 1,2,\ldots,p$ and $j = 1,2,\ldots,N$. The elements $a_1,a_2,\ldots,a_p$ are positive constants expressing the relative importance of the different increments, ${}_o\hat{\delta}_b$, and w is a positive constant expressing the relative importance of the approximating sum of squares ϕ^*. Let $\theta(w)$ denote that point in the parameter space at which $\bar{\phi}^*$ achieves its minimum, then Levenberg determines the constants $w,a_1,\ldots,a_p$ by the following procedure:

(1) The best theoretical value of w would be determined by solving

$$\frac{d\phi(\theta(w))}{dw} = 0; \tag{2.4}$$

however, this equation is usually difficult to solve so an approximate method will be used. The approximation involves setting the following expression equal to zero:

$$\phi(\theta(w)) \doteq \phi({}_{o}\hat{\theta}) + w\left(\frac{d\phi(\theta(w))}{dw}\right)_{w=o} . \tag{2.5}$$

(2) The constants $a_1,a_2,\ldots,a_p$ are chosen so that the directional derivative of ϕ, taken at $w = 0$ along the curve $\theta = \theta(w)$, should have its minimum value, namely, the negative gradient. Levenberg shows that this criterion is satisfied when $a_1,a_2,\ldots,a_p$ are all equal.

Levenberg also demonstrates the following for his procedure:

(1) The minimization of (2.2) also diminishes the sum of squares of the approximating residuals ϕ^*.

(2) The increments given by the standard least squares solution are improved.

(3) Values of w can be found for which the sums of squares of the true residuals $\phi(\theta)$ can be reduced.

(4) The usual least squares solutions for θ correspond to the case where $w\to\infty$ and hence is a special case of this procedure.

Another iterative procedure that can be used to determine those values of θ that minimize the expression $\phi(\theta)$ is the gradient method or the method of steepest descent. This method is similar to the Gauss-Newton method in that one has a preliminary estimate of θ, denoted by ${}_{o}\hat{\theta}$, and attempts to find a new estimate for θ, say ${}_{1}\hat{\theta}$, which is better than ${}_{o}\hat{\theta}$ in the sense that $\phi({}_{1}\hat{\theta}) < \phi({}_{o}\hat{\theta})$.

The gradient or steepest descent method merely steps off from the current preliminary trial value in the direction of the negative gradient of $\phi(\theta)$. One limitation with the various gradient methods is the one of slow convergence. In order to circumvent some of the problems inherent in the gradient and Gaussian estimation techniques, Marquardt [1963] develops "a maximum neighborhood method [that], in effect, performs an optimum interpolation between the Taylor series [Gauss-Newton] method and the gradient method, the interpolation being based upon the maximum neighborhood in which the truncated Taylor series gives an adequate representation of the nonlinear model." Marquardt's method involves solving the following equation for ${}_o\hat{\delta}$:

$$({}_oF^T {}_oF + \zeta I)\, {}_o\hat{\delta} = {}_oF^T {}_oy \qquad (2.6)$$

where ${}_oF = \{{}_of_j^{(b)};\ j = 1,2,\ldots,N,\ b = 1,2,\ldots,p\}$ is an $N \times p$ matrix where ${}_of_j^{(b)} = \frac{\partial f(X_j;{}_o\hat{\theta})}{\partial \theta_b}$; ${}_o\hat{\delta} = ({}_o\hat{\delta}_b, b = 1,2,\ldots,p)$ is a $p \times 1$ vector where ${}_o\hat{\delta}_b = \theta_b - {}_o\hat{\theta}_b$; ${}_oy = (y_j - f(X_j;{}_o\theta),$ $j = 1,2,\ldots,N)$ is an $N \times 1$ vector; and $\zeta \geq 0$. The superscript T on a matrix or vector represents the transpose of the corresponding matrix or vector. An outline of Marquardt's estimation procedure is given as follows:

Let ${}_oB = {}_oF^T {}_oF$ and define the new matrix ${}_oB^*$ by

$$ {}_oB^* = \{ {}_ob^*_{rr'}, r, r' = 1,2,\ldots,p\} = \left\{ \frac{{}_ob_{rr'}}{\sqrt{{}_ob_{rr}\,{}_ob_{r'r'}}} \right\}, \quad r, r' = 1,2,\ldots,p \quad . $$

Let ${}_oG = {}_oF^T\,{}_oy$ and define ${}_oG^* = ({}_og^*_1, \ldots, {}_og^*_p) = \left(\frac{{}_og_1}{\sqrt{{}_ob_{11}}}, \ldots, \frac{{}_og_p}{\sqrt{{}_ob_{pp}}} \right)^T$. At the ω^{th} iteration solve the following equation for ${}_\omega\hat{\delta}^*$:

$$ ({}_\omega B^* + {}_\omega\zeta I)\,{}_\omega\hat{\delta}^* = {}_\omega G^*. \tag{2.7} $$

Then ${}_\omega\hat{\delta}$ is obtained by ${}_\omega\hat{\delta} = \left(\frac{{}_\omega\hat{\delta}^*_1}{\sqrt{{}_\omega b_{11}}}, \ldots, \frac{{}_\omega\hat{\delta}^*_p}{\sqrt{{}_\omega b_{pp}}} \right)^T$

A new trial vector is found by taking ${}_{\omega+1}\hat{\theta} = {}_\omega\hat{\theta} + {}_\omega\hat{\delta}$ and the procedure is continued until ${}_\omega\hat{\delta}$ becomes sufficiently small. The aim of the procedure is to minimize ϕ in the "neighborhood over which the linearized function gives an adequate representation of the nonlinear function." For large values of ζ Marquardt demonstrates that the solution ${}_o\hat{\delta}$ in equation (2.6) rotates toward the solution for this increment found by the gradient method. For $\zeta = 0$ we can show that we obtain the usual Gaussian estimate for the vector of increments. Also it can be shown that equation (2.6) is the same as $\frac{d\bar{\phi}^*}{d\,{}_o\hat{\delta}} = 0$ from Levenberg's procedure for the special case when $a_1 = a_2 = \ldots = a_p = a$ from equation (2.2). For this case ζ in equation (2.6) is equal to $\frac{a}{w}$ where w has been defined earlier. Therefore during the iterative procedure small values of $_{\omega}\zeta$ are chosen when conditions are such that the Gauss-Newton method will converge nicely, which is usually true in the later steps of iteration.

The estimation of the parameters in a nonlinear regression equation is usually initiated by reducing the nonlinear function to a type of linear function. In the iterative procedures discussed in the previous paragraphs, this was acccmplished by a Taylor series expansion of the nonlinear function about some preliminary estimate of the vector of parameters. Since many physical laws which are represented by nonlinear functions are derived from simple, mainly linear, relationships between the function and its first and second derivatives, Hartley [1948], Lipton and McGilchrist [1964], and Wiggins [1960] have attempted to replace the original nonlinear regression equations by equivalent linear finite difference equations. Then they perform a least squares estimation procedure on the set of linear equations.

As an example consider the regression function:

$$E(y) = \theta_1 - e^{-\theta_2 x} \tag{2.8}$$

which is generated by the first order differential equation

$$\frac{dE(y)}{dx} = \theta_2(\theta_1 - E(y)). \tag{2.9}$$

This differential equation is equivalent to the following difference equation

$$E(y_{j+1}) - E(y_j) = (\theta_1 - E(y_j))(1 - e^{-\theta_2 h}) \tag{2.10}$$

where $E(y_j) = \theta_1 - e^{-\theta_2 x_j}$ and $x_{j+1} - x_j = h$ for all j. In a more general form the equation may be written as:

$$E(y_{j+1}) - E(y_j) = aE(y_j) + b. \tag{2.11}$$

This is the type of difference equation considered by Hartley. From a knowledge of summation of finite differences, equation (2.11) takes the form

$$E(y_j) - E(y_o) = a \sum_{j'=o}^{j-1} E(y_{j'}) + bx_j + c \tag{2.12}$$

where c is a constant of summation. Hartley's "internal least squares approach" may formally be described as follows:

(1) The nonlinear regression equation, which is the solution of a linear differential equation, is replaced by an equivalent linear finite difference equation.

(2) The observations y_j are expressed as a linear function of the progressive sums $\sum_{j'=o}^{j-1} E(y_{j'})$ and the independent variable x_j, by forming the progressive sums on the equivalent finite difference equation and replacing $E(y_j)$ by y_j.

(3) Finally, a least squares fit is made on this last linear

equation, from which estimates of the original parameters are found.

Lipton and McGilchrist [1964] present a general method of estimation for the parameters in the multiple exponential regression function from an equivalent finite difference equation. For the multiple exponential curve

$$E(y_j) = \alpha + \sum_{i=1}^{k} \beta_i \rho_i^j \ , \ j = 0,1,\ldots,N-1, \tag{2.13}$$

they

(1) Show that the following general finite difference equation is capable of generating (2.13):

$$G_{m+k-1}E(y_{j+k}) + G_{m+k-2}E(y_{j+k-1}) +\ldots+ G_o E(y_{j-m+1}) + \alpha G = 0 \tag{2.14}$$

where $m \geq k$ and the G's are specified functions of the parameters.

(2) Substitute the observations y_{j+k} for $E(y_{j+k})$ in (2.14).

(3) Suggest estimating the parameters by minimizing either of the following expressions:

$$\sum_{j=-k}^{N-k-1} \left(G_{m+k-1}y_{j+k} + G_{m+k-2}y_{j+k-1} +\ldots+ G_o y_{j-m+1} + \alpha G \right)^2 \tag{2.15}$$

or

$$\sum_{d=-k}^{N-k-1} \left[\sum_{j=-k}^{d} \left(G_{m+k-1}y_{j+k} + G_{m+k-2}y_{j+k-1} +\ldots+ G_o y_{j-m+1} + \alpha G \right) \right]^2 . \tag{2.16}$$

They also indicate that several other nonlinear models may be represented in the form (2.14).

Wiggins [1960] also presents an estimation procedure that is based upon some of the concepts of finite differences. Moreover, his procedure is also presented for simultaneous estimation for more than one nonlinear equation. This procedure may be briefly described as follows:

(1) Let y_j, $j = 1,2,\ldots,N$, be a set of N observations.

(2) Let $E(y_j)$ be a function of the independent variable x_j, where $x_{j+1} - x_j = h$ for all j.

(3) Assume that $\frac{dE(y)}{dx}$ is expressable as a linear function of E(y) and the parameters to be estimated. Then replace $\frac{dE(y)}{dx}$ by

$$\frac{y_{j+1}-y_{j-1}}{x_{j+1}-x_{j-1}} = u_j \text{ and } E(y_j) \text{ by } y_j.$$

(4) Fit the linear equation which expresses u_j as a linear function of y_j and the parameters by least squares.

Again this procedure is applicable to any physical situation that may be described by means of a system of linear differential equations with constant coefficients.

Besides the above papers that have been concerned with the development of least squares or pseudo least squares estimation procedures, there have been some related papers that considered aspects of least squares estimation other than the estimation problem. One of these problems is concerned with measuring the amount of departure from linearity by a nonlinear regression model. This is a pertinent consideration since many of the least squares techniques approximate a nonlinear function by a truncated linear Taylor series expansion of the nonlinear function. Beale [1960] has proposed a measure of nonlinearity given by:

$$\hat{N}_\theta = \frac{p\hat{\sigma}^2 \sum_{\iota=1}^{m} \sum_{j=1}^{N} \left[f(X_j;{}_\iota\theta) - f(X_j;\hat{\theta}) - \sum_{b=1}^{p} ({}_\iota\theta_b - \hat{\theta}_b) \frac{\partial f(X_j;\hat{\theta})}{\partial \theta_b} \right]^2}{\sum_{\iota=1}^{m} \left\{ \sum_{j=1}^{N} \left[f(X_j;{}_\iota\theta) - f(X_j;\hat{\theta}) \right]^2 \right\}^2} \tag{2.17}$$

where $\hat{\sigma}^2$ is the estimate of $\sigma^2 = E(\varepsilon_j^2)$, $j = 1,2,\ldots,N$; $E(y_j) = f(X_j;\theta)$; $\hat{\theta}$ is the least squares estimate of θ; and ${}_\iota\theta$, $\iota = 1,2,\ldots,m$, represent points in the neighborhood of $\hat{\theta}$. Beale concludes that the model is "disastrously nonlinear in θ" if $\hat{N}_\theta > {}^1/F_\alpha(p,\nu)$ while the linear approximation to the nonlinear model is satisfactory if $\hat{N}_\theta < 0.01/F_\alpha(p,\nu)$ where $F_\alpha(p,\nu)$ is the upper $100\alpha\%$ point of the F distribution with (p,ν) degrees of freedom and ν is the degrees of freedom associated with the estimate of σ^2. Although Beale only speaks of $\hat{\sigma}^2$ as being an "adequate independent estimate" of σ^2 both Beale and Guttman and Meeter [1965], who examine the validity and usefulness of this measure of nonlinearity by means of numerical examples, propose and use the following estimate of σ^2: $\hat{\sigma}^2 = \sum_{j=1}^{N} (y_j - f(X_j;\hat{\theta}))^2/(N-p)$ where $\hat{\theta}$ is the least squares estimate of θ.

Other estimation procedures that are concerned with the specific problem of estimating the parameters in the class of regression models given by (1.1) are the partial totals technique discussed by Cornell [1962] and the Spearman estimation procedure as presented by Johnson and Brown [1961]. Both of these procedures have limited their consideration to special members of the class of models given by (1.1). More specifically the partial totals procedure has been limited to the case when $n = 1$ and the Spearman estimation procedure has been limited to the case $n = m = 1$ in equation (1.1). For the special case $n = m = 1$ in (1.1) there have been many estimation procedures presented for the estimation of the exponential parameter and Speckman and Cornell [1965] outline for this case some of the more familiar estimation techniques, i.e. maximum likelihood, least squares, weighted least squares, and partial totals. After evaluating the above procedures these authors conclude that the maximum likelihood and partial totalsmethods give similar results for small values of N. Since the main contribution of this research has been the generalization of the partial totals and Spearman estimation procedures to the simultaneous consideration of several equations, we will wait until Chapters 5 and 6 before giving the details of these estimation techniques. The results of Speckman and Cornell [1965], the simplicity of these two estimation schemes for the simple exponential model, and the high efficiency (88%) of the Spearman estimation procedure for the simple exponential model with binomial variation have provided a great deal of the motivation for this research.

2.3 Simultaneous estimation schemes

Although there have been some papers written on the problem of estimating the parameters in a set of regression functions simultaneously, the papers have not considered the particular situation covered by this research. Telser [1964] and Zellner [1962] have presented simultaneous estimation schemes that may be used when one is faced with a set of linear regression equations each being a function of a different set of parameters. Much of the estimation development in these papers is based on Aitken's [1934] generalized least squares for linear equations. With respect to the problem of simultaneous estimation of the parameters in a set of nonlinear regression equations, there has been very little written. Box and Draper [1965] have considered the Bayesian estimation of common parameters from several responses when the observed random variables of our regression model are assumed to follow a multivariate normal distribution. In addition, Turner, et. al. [1963] have considered this problem when the covariance matrix for the error terms is assumed to be known to within a constant multiplier. Turner's approach is based upon a Taylor series expansion of each regression equation through the linear terms about some preliminary estimates of the parameters, and then an approach similar to that used by Telser [1964] or Zellner [1962] is applied iteratively. Beauchamp and Cornell [1966] use a similar procedure for a system of nonlinear regression equations in which some of the parameters may be common to more than one of the regression equations; however, fewer assumptions are made about the covariance matrix of the error terms. The details of this simultaneous nonlinear estimation procedure are included in Chapter 4.

2.4 Models

This research was initially motivated by investigations reported in an article by Galambos and Cornell [1962] involving the use of radioactive tracers in a biological experiment. Therefore throughout this research we have attempted to formulate and generalize our simultaneous estimation procedures with the ultimate purpose in mind of applying them to the estimation of the parameters used to describe tracer experiments. Sheppard [1962] has presented one of the more complete discussions on the basic concepts of the use of tracers beginning with a discussion of the elementary principles of the tracer method. Sheppard then moves into a discussion of tracer experiments in compartmental systems and the problem of model building. Properties of these models will be derived in Chapter 3 which will be useful in the development of estimation procedures in later chapters. Sheppard also includes an extensive bibliography in his book on tracer experiments. Berman and Schoenfeld [1956] also give a good discussion of the formulation of models for tracer experiments in steady state, and consider the problem of estimating the constants that are inherent in the physical or biological experiment. Cornfield, et. al. [1960] are also concerned with the problem of model building and estimation as related to tracer experiments; however, they only consider one equation at a time in their estimation process. Berman [1961] gives a good example of the application of tracer experimentation and model building to the thyroid system. A general discussion of compartmental models will be given in this paper in Chapter 3.

III. MODELS TO BE CONSIDERED

3.1 General discussion of compartmental models

In the introduction to this research the statement was made that there exist examples where a regression model such as that given by equation (1.1) is used to describe a physical situation. Two such examples that arise in experiments concerned with the use of radioactive substances as tracer material are the mammillary and catenary systems. Although a great deal of writing has been done concerning the mathematical formulation of models for tracer experiments, two of the better discussions on this subject are those by Berman and Schoenfeld [1956] and Sheppard [1962]. The basic rule of these formulations is to consider a system within an organism as made up of a number of chemical states or sites of a physiological substance. It is assumed that there are fixed transition probabilities or turnover rates from one state or site to another and the whole system is assumed to be in steady state. Hence by introducing radioactive substances into the system we are able to study the system "in vivo" without affecting the turnover rates of the system. The system as a whole is quite complicated but an adequate model for studying such processes consists of a finite number of states or compartments with turnover rates which are proportional to the amounts of material in the compartments. It is also assumed that the tracer material mixes uniformly with its isotope and that its behavior reflects that of the unlabeled substance. The concept of dividing a biological system into a number of fixed compartments is merely an aid in analysis, since the various states or sites contain finer structure. However, the compartmental analysis does prove itself useful in understanding some of the mechanics of the system.

The mammillary and catenary systems are particular examples of compartmentalized systems in steady state and they may be formally described by means of the following definitions:

Definition 3.1: The mammillary system involves n peripheral compartments that have turnover rates with a central compartment but no turnover between the n peripheral compartments.

Definition 3.2: The catenary system involves (n+1) compartments that may be thought of as arranged in a chain-like manner where each compartment has non-zero transition rates only with the compartments adjacent to it.

For a detailed discussion of these systems one may refer to the work by Sheppard [1962].

In order to show how a regression model such as that given by equation (1.1) arises, we will derive the regression model for the general (n+1)-compartmental problem and then give the particular solutions for the mammillary and catenary systems. Although the following derivations may be found elsewhere, they are presented here for completeness. The observations will be taken at particular time values; therefore the independent quantity x_j will represent a particular point in time. Since we will want to consider the expected values of our observations as being continuous functions of time, we will denote this by writing x in place of x_j, i.e. x denoting any arbitrary time point and x_j representing a particular fixed time point. The following notation will be used:

$E(Y_i(x))$ = the expected amount of labeled material in the i^{th} compartment at time x;

τ_{rs} = the fractional amount of material in the s^{th} compartment flowing to the r^{th} compartment per unit time;

$E(\eta_i(\rho))$ = the Laplace transform of $E(Y_i(x))$

$$= \int_0^\infty E(Y_i(x))e^{-\rho x}dx.$$

From the discussion in the preceding paragraphs, the following set of differential equations is formed to describe the general (n+1)-compartmental problem:

$$\frac{dE(Y_i(x))}{dx} = -E(Y_i(x)) \sum_{\substack{r=1 \\ r\neq i}}^{n+1} \tau_{ri} + \sum_{\substack{i'=1 \\ i'\neq i}}^{n+1} \tau_{ii'}E(Y_{i'}(x)) \tag{3.1}$$

for i = 1,2,...,n+1.

From our knowledge of Laplace transforms we may write the system of equations (3.1) as follows:

$$\begin{aligned}
(\rho+\tau_{11})E(\eta_1(\rho)) - \tau_{12}\,E(\eta_2(\rho)) - \dots -\tau_{1,n+1}E(\eta_{n+1}(\rho)) &= E(y_1(0))\\
-\tau_{21}E(\eta_1(\rho)) + (\rho+\tau_{22})E(\eta_2(\rho)) - \dots -\tau_{2,n+1}E(\eta_{n+1}(\rho)) &= E(y_2(0))\\
\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad & \quad \vdots\\
-\tau_{n+1,1}E(\eta_1(\rho)) - \tau_{n+1,2}E(\eta_2(\rho)) - \dots + (\rho+\tau_{n+1,n+1})E(\eta_{n+1}(\rho)) &= E(y_{n+1}(0))
\end{aligned} \tag{3.2}$$

where $\tau_{ss} = \sum_{\substack{r=1\\ r\neq s}}^{n+1} \tau_{rs}$. Using matrix notation we may write (3.2) as:

$$(\rho I + \tau)E(\eta(\rho)) = E(Y(0)) \tag{3.3}$$

where $E(Y(0)) = \Big(E(Y_1(0)),\dots,E(Y_{n+1}(0))\Big)^T$, $E(\eta(\rho)) =$

$$= \Big(E(\eta_1(\rho)),\dots,E(\eta_{n+1}(\rho))\Big)^T,$$

τ is the $(n+1)\times(n+1)$ matrix of coefficients of (3.2) with $\rho = 0$, and I is an $(n+1)\times(n+1)$ identity matrix. From matrix algebra we know that

$$(\rho I + \tau)^{-1} = \frac{\{\Delta_{rs}(\rho),\ r,s=1,2,\dots,n+1\}^T}{|\rho I + \tau|} \tag{3.4}$$

where $\Delta_{rs}(\rho)$ is the r^{th} row and s^{th} column cofactor of $(\rho I + \tau)$ and $|\rho I + \tau|$ represents the determinant of $(\rho I + \tau)$. From these results we may write

$$E(\eta_i(\rho)) = \sum_{i'=1}^{n+1} \frac{\Delta_{i'i}(\rho)}{|\rho I+\tau|}\,E(Y_{i'}(0)). \tag{3.5}$$

Denote the roots of $|\rho I + \tau| = 0$ by $-\lambda_1,-\lambda_2,\dots,-\lambda_{n+1}$ since this is a polynomial of $(n+1)^{st}$ degree in ρ. If the λ's are distinct then

$$E(\eta_i(\rho)) = \sum_{k=1}^{n+1} \frac{\alpha_{ik}}{(\rho+\lambda_k)} \tag{3.5a}$$

where

$$\alpha_{ik} = [E(\eta_i(\rho))(\rho+\lambda_k)]_{\rho = -\lambda_k} \ . \tag{3.6}$$

Finding the inverse Laplace transform of $E(\eta_i(\rho))$, the solutions for $E(Y_i(x))$ are given by:

$$E(Y_i(x)) = \sum_{k=1}^{n+1} \alpha_{ik} e^{-\lambda_k x}, \ i = 1,2,\ldots,n+1. \tag{3.7}$$

We now combine the above results into the following theorem:

Theorem 3.1: For the general (n+1)-compartmental tracer experiment where transfer of labeled material is allowed between any two compartments, the solutions for the expected amount of labeled material in the i^{th} compartment at time x is given by (3.7) if the roots of the equation $|\rho I + \tau| = 0$, given by $-\lambda_1,-\lambda_2,\ldots,-\lambda_{n+1}$, are distinct. The constants α_{ik} are given by (3.6).

From the definition of the characteristic roots of τ, we note that $\lambda_1,\lambda_2,\ldots,\lambda_{n+1}$ are equal to the characteristic roots of τ. The assumption that the λ's are distinct was made in order to arrive at a unique regression model. Otherwise, from a knowledge of Laplace transforms we would have a different model for each number of multiple roots and also for each different multiplicity of each multiple root. Therefore all of our discussion pertaining to regression models of the form given by (1.1) or (3.7) will be made under the assumption that the λ_k's are distinct whether this assumption is explicitly stated or not.

In addition, the estimation procedures developed in Chapters 5 and 6 have been derived specifically for models of the form given by (1.1) and (3.7). If it does happen that multiple characteristic roots of τ exist, the compartmental model may be modified by a procedure similar to that which is contained in the discussion pertaining to Figure 3.2 in order to eliminate the possibility of multiple roots.

Throughout the above discussion we have assumed that the τ_{rs}'s are all ≥ 0, and Berman and Schoenfeld [1956] have shown that this restriction, along with $\tau_{rr} = \sum_{\substack{s=1 \\ s \neq r}}^{n+1} \tau_{sr}$, is enough to ensure that the λ_k's have positive real parts, and that pure imaginary characteristic roots of τ are impossible.

For many of the tracer experiments that one will be faced with, a fixed amount of tracer material will be present. Therefore there are only n independent regression equations since $\sum_{i=1}^{n+1} E(Y_i(x))$ must be equal to the fixed amount of labeled material present for all values of x. We are now able to prove the following theorem which will be useful in the presentation of the estimation procedures in Chapters 5 and 6:

Theorem 3.2: For the mammillary and catenary systems, the number of exponential terms in each regression equation will be equal to the number of independent regression equations for the case when a fixed amount of tracer material is injected into the system.

Proof: For the general (n+1)-compartment mammillary system, the matrix τ will take the form:

$$\tau = \begin{pmatrix} \tau_{11} & 0 & 0 & \cdots & -\tau_{1,n+1} \\ 0 & \tau_{22} & 0 & \cdots & -\tau_{2,n+1} \\ 0 & 0 & \tau_{33} & \cdots & -\tau_{3,n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ -\tau_{n+1,1} & -\tau_{n+1,2} & -\tau_{n+1,3} & \cdots & \tau_{n+1,n+1} \end{pmatrix}$$

$$= \begin{pmatrix} \tau_{n+1,1} & 0 & 0 & \cdots & -\tau_{1,n+1} \\ 0 & \tau_{n+1,2} & 0 & \cdots & -\tau_{2,n+1} \\ 0 & 0 & \tau_{n+1,3} & \cdots & -\tau_{3,n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ -\tau_{n+1,1} & -\tau_{n+1,2} & -\tau_{n+1,3} & \cdots & \tau_{1,n+1}+\cdots+\tau_{n,n+1} \end{pmatrix}. \quad (3.8a)$$

By a number of elementary row and column operations on the matrix τ, it may be reduced to the following equivalent matrix:

$$\begin{pmatrix} \tau_{n+1,1} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \tau_{n+1,2} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \tau_{n+1,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \tau_{n+1,n} & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \quad (3.8b)$$

which has rank n and this rank is equal to the rank of τ. From equation (3.7) we note that the number of exponential terms in each regression equation of our model is determined by the number of nonzero characteristic roots of the matrix τ. The number of nonzero characteristic roots of τ is equal to the rank of τ (see Hohn [1964], page 280) and from (3.8b) we note that the rank of τ is equal to n. Since we are assuming that a fixed amount of tracer material is present in the system, n is also the number of independent regression equations and our conclusion follows for the general mammillary system.

For the general (n+1)-compartment catenary system the matrix τ is given by:

$$\begin{pmatrix} \tau_{21} & -\tau_{12} & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\tau_{21} & \tau_{12}+\tau_{32} & -\tau_{23} & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\tau_{32} & \tau_{23}+\tau_{43} & -\tau_{34} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\tau_{n,n-1} & \tau_{n-1,n}+\tau_{n+1,n} & -\tau_{n,n+1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\tau_{n+1,n} & \tau_{n,n+1} \end{pmatrix}. \tag{3.9a}$$

By a number of elementary row and column operations τ can be reduced to the following equivalent diagonal matrix of rank n:

$$\begin{pmatrix} \tau_{21} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \tau_{32} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \tau_{43} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \tau_{n+1,n} & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} . \tag{3.9b}$$

Again we use the fact that the number of nonzero characteristic roots of τ is equal to the rank of τ, i.e., n. Hence we conclude that the regression model used to describe the general $(n+1)$-compartment catenary system when a fixed amount of tracer material is present will be a set of n independent regression equations each being a linear combination of the same n exponential terms. This completes the proof of the theorem.

From Theorem 3.1 we can now state the following obvious corollary without proof:

Corollary 3.3: The regression models used to describe the general $(n+1)$-compartment mammillary and catenary systems when a fixed amount of tracer material is present in these systems, are members of the class of regression models given by equation (1.1) with $m = n$.

In addition, for the case when the amount of tracer material in the system is known and fixed, we can divide each of the equations in (3.7) by this constant and have a system of regression equations still of the form given by (1.1) in terms of a new quantity which represents the proportion of labeled material in the compartments at a time x. Besides the two very general classes of models contained within the mammillary and catenary systems discussed above, there also exist other compartmental models that are neither mammillary nor catenary in nature, but they still give rise to a regression model contained within the class of models of equation (1.1) for $m = n$. One such example is given in Figure 3.1 where the four numbered boxes represent certain chemical states or physiological sites, the arrows represent the direction of certain changes or transitions that take place, and the τ_{rs}'s represent the nonzero transition or turnover rates. Figure 3.1 is drawn only as a visual aid and is not meant to be an exact representation of the physical situation. From this figure we note that this example is neither mammillary nor catenary in nature. For this example the matrix τ is given by:

$$\tau = \begin{pmatrix} \tau_{21}+\tau_{41} & -\tau_{12} & -\tau_{14} & 0 \\ -\tau_{21} & \tau_{12}+\tau_{32} & 0 & 0 \\ 0 & -\tau_{32} & \tau_{13} & 0 \\ -\tau_{41} & 0 & 0 & 0 \end{pmatrix} . \tag{3.10}$$

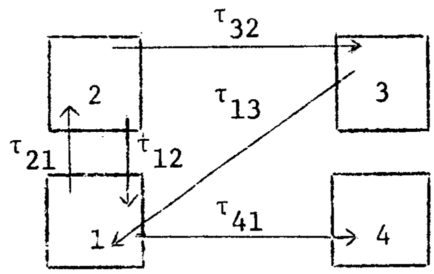

Fig. 3.1--Compartmental model which is neither mammillary nor catenary in nature.

which can be shown to have rank three. The number of characteristic roots of τ is also equal to three and hence the number of exponential terms in each regression equation is equal to three. Therefore if a fixed amount of tracer material is introduced into the system of Figure 3.1, then there are three independent regression equations and each equation will involve a linear combination of three exponential terms. Again we see that this example is a member of the class of regression models of equation (1.1).

In some practical situations it may happen that even with a fixed amount of tracer material present in the compartmental system that it is impossible to determine the amount of tracer material in each compartment. One possible way to circumvent this dilemma is to propose a simplified model where some of the unobservable compartments have been replaced by a single compartment. One possible example is given in Figure 3.2(a) where we assume that the observations made on compartments 2 and 3 are not easily resolved. For this model the matrix τ has rank 2, therefore

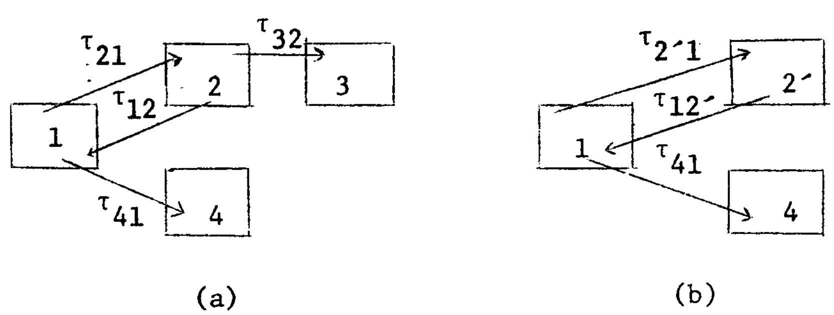

Fig. 3.2--Simplification of a compartmental model

each of the four regression equations will be a linear combination of the same two exponential terms. If neither compartment 2 nor 3 is observable then this more complex model might be replaced by the simpler model given in Figure 3.2(b). Compartment 2′ represents the combining together of the original compartments numbered 2 and 3 into a single compartment. By means of this modification, for the case when a fixed amount of labeled or tracer material is present, our regression model may be represented by a system of two independent regression equations each being a linear combination of two exponential terms. The above discussed modifications are similar to the ones proposed by Berman [1961] who was concerned with the application of tracer experiments to the thyroid system.

Another auxiliary problem that should be considered is the problem that pertains to relating the coefficients and exponential parameters of the compartmental models, which are considered here, to the turnover rates or transition probabilities of the original system. If we are considering a general (n+1)-compartment system and if we assume that nothing is known about some of the τ_{rs}'s, then we may use the following equation derived by Berman and Schoenfeld [1956] to relate the coefficients and exponential parameters of equation (3.7) to the turnover rates or transition probabilities:

$$\tau = \alpha\lambda\alpha^{-1} \tag{3.11}$$

where α is an $(n+1) \times (n+1)$ matrix of the α_{ik} from (3.7) and λ is an $(n+1) \times (n+1)$ diagonal matrix with diagonal elements λ_k. The estimates of the elements of τ would then be found by substituting the estimates of the coefficients and exponential parameters into α and λ, respectively. If some of the τ_{rs}'s are known, such as the case for the mammillary and catenary systems when some of the τ_{rs}'s are equal to zero, then it is easily seen that certain restrictions are imposed upon the other parameters of our regression model. In Section III of the paper by Berman and Schoenfeld they propose for this situation a method of estimating the nonzero elements of τ by means of transformations of the matrix α in (3.11) which preserve the constraints on the elements of α. This approach then determines a whole class of models which satisfy the initial constraints. Another approach would be to substitute the estimated regression equations into the original differential equations giving us a system of linear equations in the τ_{rs}'s after equating the coefficients of like exponential terms to each other. Since many observations are made on each equation, usually we would have more equations linear in the τ_{rs}'s than there are nonzero τ_{rs}'s. Therefore the usual least squares procedure on the complete set of linear equations could be used to determine estimates of the nonzero τ_{rs}'s. Although the main concern of the following chapters will be the estimation of parameters in the regression models of equations (1.1) and (3.7), we see that it is possible to obtain estimates of the turnover rates and transition probabilities from the estimates of the coefficients and exponential parameters of these equations.

3.2 Distributional assumptions concerning the random variables ε_{ij}

So far in our discussion no assumptions have been made about the random variables ε_{ij} that appear in equations (1.1) and (1.2). In this section we will state the general distributional assumptions to be used for all of the estimation techniques to be developed and then we will list the various modifications that will be used for particular estimation schemes. The following notation will be needed before the assumptions concerning the ε_{ij} are stated:

1) $\varepsilon_{i*} = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{iN})^T$ for $i = 1,2,\ldots,n$;

2) $\varepsilon_{*j} = (\varepsilon_{1j}, \varepsilon_{2j}, \ldots, \varepsilon_{nj})^T$ for $j = 1,2,\ldots,N$;

3) $\varepsilon_{**} = (\varepsilon_{1*}^T, \varepsilon_{2*}^T, \ldots, \varepsilon_{n*}^T)^T$ which will be an $nN \times 1$ vector;

4) $\varepsilon \ = (\varepsilon_{*1}^T, \varepsilon_{*2}^T, \ldots, \varepsilon_{*n}^T)^T$ which also will be an $nN \times 1$ vector.

For each value of i and j we assume that $E(\varepsilon_{ij}) = 0$ and for each value of i we assume that the N components of ε_{i*} are generated by independent random drawings from a given distribution with $E(\varepsilon_{i*}\varepsilon_{i*}^T) = D_{ii}$ and $E(\varepsilon_{i*}\varepsilon_{i'*}^T) = D_{ii'}$, where D_{ii} is a diagonal matrix with positive diagonal elements and $D_{ii'}$ is a diagonal matrix with nonzero diagonal elements for $i \neq i'$ and $i,i' = 1,2,\ldots,n$. Furthermore, we assume that the set of N random vectors ε_{*j} are independent vectors drawn from a given multivariate distribution. These assumptions allow for correlation between observations on different equations when they are made for the same value of the independent variable X_j. In Chapters 5 and 6 as well as in the examples considered in this research, the dimension of X is 1x1 and X is the independent variable time. Therefore for this case these assumptions would mean that we allow for correlation between observations made at the same point in time.

In Chapter 4 where a generalized least squares procedure is developed and in Chapter 5 where a generalized partial totals estimation procedure is developed, much of the work will be done under the assumption that $D_{ii} = \sigma_{ii} I$ and $D_{ii'} = \sigma_{ii'} I$ for $i, i' = 1, 2, \ldots, n$ and $i \neq i'$ where $0 < \sigma_{ii} < \infty$, $-\infty < \sigma_{ii'} < \infty$, and I is an $N \times N$ identity matrix. This will also imply that $E(\varepsilon\varepsilon^T) = I \otimes \sigma_{**}$ where I is an $N \times N$ identity matrix, $\otimes$ represents the Kronecker or direct product of two square matrices, and $\sigma_{**} = \{\sigma_{ii'}; i, i' = 1, 2, \ldots, n\}$. However, a discussion will be given to a consideration of the modifications that arise when the general distributional assumptions given above are satisfied. Also in Chapter 4 we will alter some of the above restrictions to allow for complete independence among the observations and, in addition, we will allow $E(\varepsilon\varepsilon^T) = \sigma I$ where $0 < \sigma < \infty$ and I is an $nN \times nN$ identity matrix. The altering of the assumptions is done in order to investigate the simplifications that arise. In Chapter 6 where a generalization to the Spearman estimation procedure is presented, we develop the procedure under the general distributional assumptions given earlier.

So far we have not discussed the specific form of the distribution function of the random variables ε_{ij}, and for the development of the estimation procedures no specific form is required although it is assumed that $E(\varepsilon_{ij}) = 0$ for all i and j. However, with the above stated assumptions we are able to investigate some properties of our estimators in the following cases:

1) The vectors ε_{*j}, $j = 1,2,\ldots,N$, each have a multivariate normal distribution.

2) The vectors ε_{*j}, $j = 1,2,\ldots,N$, are each distributed according to the multinomial distribution.

It is quite obvious that we have not considered all possible distributions for the random variables ε_{ij}. However, by investigating the estimation procedures under the normality and multinomial assumptions we have considered distributions of practical importance. Moreover, under certain conditions realized in practice, distributions such as the Poisson, binomial, and multinomial tend to a normal distribution. Therefore a normal distribution may be a very good approximation even when it is not the true distribution of the random variables in our model.

For example, when we are measuring the proportion of radioactive tracer substance present at a site at a particular time, these measurements will usually be a ratio of random variables with the denominator related to the initial count and the numerator related to the count at the particular time of observation. If we could think of the denominator as being a constant and the numerator as being a Poisson or truncated Poisson random variable, then the distribution of the ratio could be approximated by a normal distribution when the mean of the Poisson random variable in the numerator and the denominator relative to the mean of the numerator are both large. This is true since a Poisson distribution can be closely approximated by a normal distribution when the mean of the Poisson distribution is large. The last requirement concerning the relative magnitude of the denominator and the mean of the numerator will mean that the probability of an observation being greater than one is small.

For those cases where our observations might be considered as a ratio of two random variables, a great deal of unwieldiness arises in the determination of the exact distribution of such a ratio as can be seen by referring to the work by Curtiss [1941], Donahue [1964], Fieller [1932], Geary [1930], Gurland [1948], and Merrill [1928]. However, there are some results that would appear to justify the use of a normal approximation to the distribution of our observations. For example, in radioactive tracer experiments various adjustments are made on the observations to account for the radioactive decay and this in turn will lead to non-integral values. More specifically, let an observation y be represented by the ratio y_1/y_2, where the distributions of y_1 and y_2 are given by $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. The correlation coefficient of y_1 and y_2 will be denoted by ρ. Merrill [1928], through the use of graphs and tables, demonstrates the approximate normal distribution of y under the following conditions:

1) The coefficient of variation (c.v.) of y_2 is small.

2) The quantity $\left(\dfrac{\sigma_2}{\mu_2} - \rho\dfrac{\sigma_1}{\mu_1}\right) \Big/ \left(\dfrac{\sigma_2^2}{\mu_2^2} - 2\rho\dfrac{\sigma_1\sigma_2}{\mu_1\mu_2} + \dfrac{\sigma_1^2}{\mu_1^2}\right)^{\frac{1}{2}}$ is small.

The first condition would imply that the standard deviation of y_2 is much smaller than the mean of y_2, and this condition is obviously satisfied for the example considered earlier when y_2 is a Poisson random variable with a large mean. It can be shown that the second condition could also be satisfied when ρ is small and the ratio $\frac{\text{c.v. of } y_2}{\text{c.v. of } y_1}$ is small. For those situations where our observations denote the proportion of radioactive tracer present at a site at a particular time, then the above conditions could be satisfied if we thought of our observation as being a ratio of independent Poisson random variables each with a large mean such that the mean of the denominator is much larger than the mean of the numerator, which will mean that ratios greater than one are unlikely. Hence the results of Merrill confirm the reasonableness of a normality assumption in many practical situations.

Since we had mentioned earlier that we would consider the case where our observed random variables would follow a multinomial distribution, now we want to demonstrate how such a situation could arise in an experimental situation. Consider the case where we have n independent Poisson random variables $Z_1, Z_2, \ldots, Z_n$ with parameters $\kappa_1, \kappa_2, \ldots, \kappa_n$, respectively. Let $z_1, z_2, \ldots, z_n$ represent the values that our respective n independent Poisson random variables take on, then for the situation where we assume that $z_1+z_2+\ldots+z_n$ remains fixed, the conditional distribution of $Z_1, Z_2, \ldots, Z_n$ follows the multinomial distribution given by:

$$\frac{(z_1+z_2+\ldots+z_n)!}{z_1!z_2!\ldots z_n!} \quad \frac{\kappa_1^{z_1}}{\kappa^{z_1}} \; \frac{\kappa_2^{z_2}}{\kappa^{z_2}} \ldots \frac{\kappa_n^{z_n}}{\kappa^{z_n}}$$

where $\kappa = \kappa_1+\kappa_2+\ldots+\kappa_n$. The above result can be shown by a direct generalization of the results given by Birnbaum [1954] and Bross [1954]. Therefore for the case when the sum of our observed independent Poisson random variables remains constant, the joint conditional distribution of these random variables is given by the appropriate multinomial distribution. As $z_1+z_2+\ldots+z_n$ becomes large, the multinomial distribution tends to an (n-1)-dimension multivariate normal distribution, hence making not only the multinomial but the normal model a reasonable approximation to the actual situation.

IV. GENERALIZED LEAST SQUARES ESTIMATION

4.1 Introduction

In this chapter a generalized least squares estimation procedure, which will be used for the estimation of the parameters in equation (1.2), will be developed and evaluated. At the present time, most of the least squares methods of estimation in nonlinear regression equations appear to have been restricted to the situation where only one regression equation is present, as the literature cited in Chapter 2 demonstrates. In addition to the articles presented in Chapter 2, one might refer to the articles by Hartley and Booker [1965], Stevens [1951], and Turner et. al. [1961] for a demonstration of this fact. With respect to estimation procedures that attempt to estimate the parameters present in a set of regression equations simultaneously, Turner et. al. [1963] present a generalized least squares estimation procedure for a set of nonlinear regression equations when the covariance matrix of the ε_{ij} terms of (1.2) is assumed to be known to within a constant multiplier. However it appears that most of the recent work has been done with a set of linear regression equations, as a reference to the articles by Telser [1964] and Zellner [1962] can show. Hence after outlining in Section 4.2 the nonlinear least squares technique for the case when n = 1 from equation (1.2), we will develop in this chapter a generalized least squares estimation procedure to be used in estimating a set of parameters simultaneously in a set of nonlinear regression equations. The results given in this chapter except for a portion of Section 4.4 are also presented in the article by Beauchamp and Cornell [1966], which was written during the time when research was being carried out for this dissertation. For the sake of completeness we will repeat these results here.

4.2 Single equation least squares estimation

Since some of the techniques used in single equation least squares estimation will be carried over into our generalized least squares procedure discussed in Section 4.3, we will now outline an iterative least squares estimation technique as we indicated in Section 2.2. Since we will present the case here when $n = 1$ from equation (1.2) and for this case $i = 1$ only, the subscript i will be dropped in order to simplify the notation.

Let y_j represent a value that the random variable Y_j has taken on for the particular value of the independent variable X_j and let $y_1, y_2, \ldots, y_N$ be a set of N independent observations drawn at random from a given population, so that $E(Y_j) = f(X_j;\theta)$ where f is a continuous differentiable function, X_j is an $h \times 1$ vector of independent quantities assumed to be known, and θ is a $p \times 1$ vector of constant parameters to be estimated. The estimation of the elements of the vector θ by least squares techniques involves the minimization of the expression:

$$\phi(\theta) = \sum_{j=1}^{N} \left(y_j - f(X_j;\theta) \right)^2 . \tag{4.1}$$

If θ_b, $b = 1,2,\ldots,p$, represents an element of the vector θ, then in order to determine the value of θ that minimizes $\phi(\theta)$, we must solve the following set of equations:

$$\frac{\partial\phi(\theta)}{\partial\theta_b} = 0, \; b = 1,2,\ldots,p. \tag{4.2}$$

If f is linear in the elements of θ, then (4.2) is a set of linear equations in the elements of θ and may be easily solved for the least squares estimator of θ, denoted by $\hat{\theta}$. For the case when f is a nonlinear function, the iterative technique presented here, which will be referred to as the Gaussian iterative technique, involves the expansion of the function $f(X_j;\theta)$ in a Taylor series about some preliminary value of θ, say ${}_o\hat{\theta}$, and truncating after the linear terms in $(\theta_b - {}_o\hat{\theta}_b)$. By using this approximation we are minimizing the expression:

$$\phi^* = \sum_{j=1}^{N} \left(y_j - {}_of_j - {}_o\hat{\delta}_1 \, {}_of_j^{(1)} - \ldots - {}_o\hat{\delta}_p \, {}_of_j^{(p)} \right)^2 \tag{4.3}$$

where ${}_o\hat{\delta}_b = (\theta_b - {}_o\hat{\theta}_b)$, ${}_of_j = f(X_j;{}_o\hat{\theta})$, ${}_of_j^{(b)} = \frac{\partial f}{\partial\theta_b}(X_j;{}_o\hat{\theta})$ for

$b = 1,2,\ldots,p$ and $j = 1,2,\ldots,N$. The Gaussian method would solve the new set of equations:

$$\frac{\partial\phi^*}{\partial\, {}_o\hat{\delta}_b} = 0, \; b = 1,2,\ldots,p, \tag{4.4}$$

for ${}_o\hat{\delta}_b$'s, and then find a new estimate of the θ_b's by taking ${}_1\hat{\theta}_b = {}_o\hat{\theta}_b + {}_o\hat{\delta}_b$ for $b = 1,2,\ldots,p$. The new vector ${}_1\hat{\theta}$ would be substituted in the place of ${}_o\hat{\theta}$ and a new set of increments ${}_1\hat{\delta}_b$ would be calculated. This process would be continued until the increments become sufficiently small.

Hartley [1961] considers the problem of convergence, which we mentioned in Section 2.2, and presents some assumptions and modifications to the Gauss-Newton estimation procedure. This modified Gauss-Newton method, which has the merit of guaranteed convergence under the assumptions to be stated below, may be briefly described as follows:

1) The first step of the modified procedure involves the determination of the vector ${}_1\hat{\delta}$ by the usual Gaussian method described previously. However in the place of the vector ${}_1\hat{\theta} = {}_o\hat{\theta} + {}_1\hat{\delta}$, Hartley uses the vector ${}_o\hat{\theta} + v_1\hat{\delta}$ where v is a scalar and $0 \leq v \leq 1$.

2) The vector ${}_o\hat{\theta} + v\,{}_1\hat{\delta}$ is substituted into $\phi(\theta)$ giving us

$$\sum_{j=1}^{N}\left(y_j - f(X_j;{}_o\hat{\theta} + v\,{}_1\hat{\delta})\right)^2 , \tag{4.5}$$

which is considered as a function of v and minimized over the range from 0 to 1, giving the value of v denoted by v_{min}. Hartley suggests the following method to approximate the value v_{min} : (i) Calculate $\phi({}_o\hat{\theta})$, $\phi({}_o\hat{\theta} + \frac{1}{2}\,{}_1\hat{\delta})$, and $\phi({}_o\hat{\theta} + {}_1\hat{\delta})$; (ii) determine the value of v for which the parabola through these three points attains its minimum; and (iii) denote this value of v by v^*_{min} and take this as an approximation to v_{min}. This parabola may be found by using the LeGrange interpolation formula.

3) The vector ${}_{o}\hat{\theta} + v_{min}\, {}_{1}\hat{\delta}$, or in most practical situations the vector ${}_{o}\hat{\theta} + v^{*}_{min}\, {}_{1}\hat{\delta}$, is substituted in the place of ${}_{o}\hat{\theta}$ and the above procedure is repeated until the vector of increments is sufficiently small. It should be noted that the desirable properties of this estimation procedure are given in terms of v_{min} instead of its approximation v^{*}_{min}.

Sufficient conditions for the convergence of the estimators found by this modified procedure to the solution of equations (4.2) using v_{min} as defined above are given as follows:

1) The first and second derivatives of $f(X;\theta)$ with respect to the elements of θ are continuous functions of the elements of θ for all X.

2) For any non-trivial set of u_b, $b = 1,2,\ldots,p$, with $\sum_{b=1}^{p} u_b^2 > 0$,

$$\sum_{j=1}^{N} \left(\sum_{b=1}^{p} u_b\, f^{(b)}(X_j;\theta) \right)^2 > 0 \qquad (4.6)$$

for the observed vectors X_j and for all θ in a bounded convex set S of the parameter space.

3) It is possible to find a vector ${}_{\omega}\hat{\theta}$ in the interior of S such that $\phi({}_{o}\hat{\theta}) < \bar{\phi}$, where $\bar{\phi} = \liminf_{\tilde{S}} \phi(\theta)$ and $\tilde{S}$ is the complement of S.

The above three assumptions on the function $f(X;\theta)$ might appear to be restrictive assumptions as stated. However, the first assumption merely allows us to define the set of equations as given by (4.3) and (4.4). The second assumption allows us to determine the solutions to equations (4.4) and is equivalent to the full rank criterion in a linear least squares problem. For a regression model specified as a linear combination of exponentials the first assumption is obviously true, and if the exponential parameters are distinct then the second assumption is obviously true. As has been pointed out by Hartley in his article, the third assumption is particularly difficult to verify if the surface represented by $\phi(\theta)$ has numerous local minima and/or maxima and/or saddle points. For this process to converge to the absolute minimum of $\phi(\theta)$, the third assumption states that we must begin our estimation procedure in a region that contains the absolute minima and no local minima. This may be difficult to verify, but in some situations it is possible to search the parameter space at a wide grid in order to locate an initial estimate ${}_o\hat{\theta}$ in the region S. However, if the parameter space is unbounded or of high dimensionality then the grid search may be unfeasible; and, in addition, the grid search implies further assumptions about the smoothness of ϕ, hence of f.

It should be noted that for the regression model given by (1.2) with n = 1 that the least squares estimator for θ is equivalent to the maximum likelihood estimator, when it is assumed that the random variables ε_j are independent with a $N(0,\sigma^2)$ distribution. This can be easily seen since maximizing the likelihood function

$$L = \prod_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\,(y_j - f(X_j;\theta))^2\right) \tag{4.7}$$

is equivalent to minimizing the expression

$$\phi(\theta) = \sum_{j=1}^{N} \left(y_j - f(X_j;\theta)\right)^2 . \tag{4.8}$$

4.3 Presentation and evaluation of estimation procedure

Referring back to equation (1.2) we note that there are nN responses y_{ij}, i.e. N independent vectors $y_{*j} = (y_{1j}, y_{2j}, \ldots, y_{nj})^T$ of observations for each value of the fixed input vector X_j. The elements of X_j are assumed to be known. Also from equation (1.2) we note that we want the same set of fixed input vectors for each regression equation or equivalently for each value of i. When X is a scalar and represents the independent variable time, this restriction means that we observe each of the n equations at the same N time values. This restriction will be relaxed in a later section of this chapter.

Before presenting the actual steps of this estimation procedure we will define the following notation:

$$y_{*j} = (y_{1j}, y_{2j}, \ldots, y_{nj})^T,\ y = (y_{*1}^T\ y_{*2}^T\ \ldots\ y_{*N}^T)^T ,$$

$$f_{*j} = \left(f_1(X_j;\theta)\ f_2(X_j;\theta)\ \ldots\ f_n(X_j;\theta)\right)^T ,$$

$$f = (f_{*1}^T\ f_{*2}^T\ \ldots\ f_{*N}^T)^T \text{ , where } X_j = (x_{j1}, x_{j2}, \ldots, x_{jm})^T .$$

With the above notation we may write equation (1.2) in the following matrix form:

$$y = f + \varepsilon \tag{4.9}$$

where ε has been defined earlier in Section 3.2. Generalizing from equation (2.1), we now want to determine a vector $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_p)^T$ which will minimize

$$\Phi = (y-f)^T \, \Omega^{-1}(y-f) \tag{4.10}$$

when evaluated at $\theta = \hat{\theta}$ for $\Omega = E(\varepsilon\varepsilon^T) = I \otimes \sigma_{**}$ as defined in Section 3.2. The generalized least squares estimation procedure to be presented will be given first for the case when Ω is assumed to be known and then it will be extended to the case when Ω is unknown.

The first step in the estimation procedure will be to expand each one of the n regression equations in a Taylor series through the linear terms about a vector of preliminary estimates of the elements of θ, say

$${}_o\hat{\theta} = ({}_o\hat{\theta}_1, {}_o\hat{\theta}_2, \ldots, {}_o\hat{\theta}_p)^T \tag{4.11}$$

to give

$$f \doteq {}_of + {}_of' \, {}_o\hat{\delta} \tag{4.12}$$

where in addition to the definitions already given we have

$$ {}_o f = ({}_o f_{*1}^T \; {}_o f_{*2}^T \cdots {}_o f_{*N}^T)^T , $$

$$ {}_o f_{*j} = \left(f_1(X_j;{}_o\hat{\theta}) \; f_2(X_j;{}_o\hat{\theta}) \cdots f_n(X_j;{}_o\hat{\theta}) \right)^T $$

for $j = 1,2,\ldots,N$,

$$ {}_o f' = \left(({}_o f_1')^T \; ({}_o f_2')^T \cdots ({}_o f_n')^T \right)^T , $$

$$ {}_o f_j' = \left\{ \frac{\partial f_i({}_o\hat{\theta};X_j)}{\partial \theta_b} \right. , \text{ row corresponds to } i = 1,2,\ldots,n; $$

column to $b = 1,2,\ldots,p \Big\}$, and ${}_o\hat{\delta} = \theta - {}_o\hat{\theta}$. For the case when Ω is assumed to be known we can use Hartley's modified Gauss-Newton procedure to calculate a vector of least squares estimates for the elements of θ in our nonlinear model. In order to keep from any ambiquity arising, the generalized form of the assumptions which were originally given in Section 4.2 are stated:

1) The first and second derivatives of any $f_i(X;\theta)$, $i = 1,2,\ldots,n$, with respect to the θ_b, $b = 1,2,\ldots,p$, are assumed to be continuous functions of the elements of θ for all vectors $X_j, j = 1,2,\ldots,N$.

2) The following inequality

$$ \sum_{j=1}^{N} \left(\sum_{b=1}^{p} u_b \frac{\partial f_i(X_j;\theta)}{\partial \theta_b} \right)^2 > 0 $$

is assumed to hold for $i = 1,2,\ldots,n$, for any non-trivial set u_b with $\sum_{b=1}^{p} u_b^2 > 0$, for the observed vectors X_j, and for all θ in a bounded convex set S of the parameter space.

3) A vector ${}_o\hat{\theta}$ exists in the interior of S such that $\Phi({}_o\hat{\theta}) < \bar{\Phi}$

where $\bar{\Phi} = \liminf_{\tilde{S}} \Phi(\theta)$ and $\tilde{S}$ is the complement of S. For the significance of these assumptions in a practical situation, one can refer to the discussion in Section 4.2.

Hartley's modified procedure starts with the usual weighted least squares estimate ${}_o\hat{\delta}$ of the increment vector $(\theta - {}_o\hat{\theta})$, where

$$ {}_o\hat{\delta} = \left[({}_of')^T \Omega^{-1} ({}_of') \right]^{-1} \left[({}_of')^T \Omega^{-1} (y - {}_of) \right]. \tag{4.13} $$

Then the vector ${}_o\hat{\theta} + v\ {}_o\hat{\delta}$ is considered where $0 \le v \le 1$, and is substituted for θ in Φ. Then Φ is considered as a function of v and this is denoted by $\Phi({}_o\hat{\theta} + v\ {}_o\hat{\delta})$. The value of v which minimizes $\Phi({}_o\hat{\theta} + v\ {}_o\hat{\delta})$ is denoted by v_{min} and the vector is ${}_o\hat{\theta} + v_{min}\ {}_o\hat{\delta}$ is substituted in the place of ${}_o\hat{\theta}$. This process is continued until the vector of increments becomes sufficiently small. The following theorem makes it possible for us to appeal to the results already proved by Hartley concerning the property of guaranteed convergence:

Theorem 4.1: Under assumptions 1 - 3 given above for the case when Ω is known, the results proved by Hartley for the single nonlinear regression model carry over to the regression model given by equation (4.9), i.e. the iterative procedure converges and provided no two stationary points of Φ yield identical values it converges to the minimum of Φ.

Proof: Since $\Omega = I \otimes \sigma_{**}$ and σ_{**} is positive definite, we know that $\Omega^{-1} = I \otimes \sigma_{**}^{-1}$ and Ω^{-1} is also positive definite (see Hohn [1964]). Therefore Ω^{-1} may be written as $U^T U$ where U is a nonsingular matrix, and Φ may be written as $Z^T Z$ where $Z = U(y-f)$, i.e. Φ may be written as a sum of squared deviations. Hence the problem has now been reduced to the same problem as that considered by Hartley in which he proved that this estimation procedure had the property of guaranteed convergence to the minimum of Φ when the assumptions are satisfied.

For the case when Ω is unknown we follow an approach similar to the one presented by Zellner [1962]. Recalling from Chapter 3 that $\Omega = I \otimes \sigma_{**}$, the problem here will be to specify estimates for the elements of the matrix σ_{**}. In order to do this we fit each of the nonlinear regression equations separately and compute a least squares estimate $\hat{\theta}^{(i)}$ of θ for the i^{th} equation, $i = 1,2,...,n$. Then the estimates of the elements of the matrix σ_{**}, which form a matrix denoted by $\hat{\sigma}_{**}$, are given by

$$\hat{\sigma}_{ii'} = \hat{\varepsilon}^T_{i*} \hat{\varepsilon}_{i'*}/N \tag{4.14}$$

where $\hat{\varepsilon}_{i*} = (\hat{\varepsilon}_{i1}\ \hat{\varepsilon}_{i2}\ \ldots\ \hat{\varepsilon}_{iN})^T$ and $\hat{\varepsilon}_{ij} = y_{ij} - f_i(X_j;\hat{\theta}^{(i)})$ for $i,i' = 1,2,...,n$ and $j = 1,2,...,N$. The following lemma demonstrates that $\hat{\sigma}_{ii'}$ is a consistent estimator of $\sigma_{ii'}$:

Lemma 4.2: Under the assumptions of Hartley's modified Gauss-Newton procedure (or any other procedure that will lead to an estimate of θ that minimizes $\sum_{j=1}^{N} \left(y_{ij}-f_i(X_j;\theta)\right)^2$ for any i) and the assumption of normality of the distribution of the vectors ε_{i*}, the estimators $\hat{\sigma}_{ii'}$ are consistent estimators of $\sigma_{ii'}$ where $i,i' = 1,2,...,n$.

Proof: From Theorem 4.1 $\hat{\theta}^{(i)}$ is a least squares estimate of θ, and since ε_{i*} is normally distributed, $\hat{\theta}^{(i)}$ is the maximum likelihood estimator of θ for the i^{th} equation and therefore it is a consistent estimator of θ. This last statement can be demonstrated by showing that the normal density function satisfies the three conditions given by Cramér ([1946], page 500) concerning the asymptotic properties of maximum likelihood estimators. In the demonstration of these sufficient conditions we use the fact that the vectors ε_{i*} are independent and identically distributed. Next consider

$$\hat{\sigma}_{ii'} = \frac{1}{N}\sum_{j=1}^{N}\left[(y_{ij}-f_i(X_j;\theta)) + (f_i(X_j;\theta) - f_i(X_j;\hat{\theta}^{(i)}))\right]$$
$$\cdot\left[(y_{i'j}-f_{i'}(X_j;\theta)) + (f_{i'}(X_j;\theta) - f_{i'}(X_j;\hat{\theta}^{(i')}))\right].$$

Using Theorem 5 and Example 4.3 from Pratt [1959] along with Khintchine's theorem (see Cramér [1946], page 254) we can show that $\hat{\sigma}_{ii'}$ converges in probability to $\sigma_{ii'}$. Hence from the definition of a consistent estimator, $\hat{\sigma}_{ii'}$ is a consistent estimator of $\sigma_{ii'}$.

The following theorem demonstrates some of the desirable asymptotic results when the previously defined estimators for the elements of Ω are used for the case when Ω is unknown.

Theorem 4.3: Under the assumptions of Lemma 4.2 the estimator of θ found by minimizing $\hat{\Phi} = (y-f)^T\,\hat{\Omega}^{-1}(y-f)$ will converge in probability to the estimator found by minimizing Φ and will be a consistent estimator of θ.

Proof: Let $\hat{\sigma}_{**} = \left\{\hat{\sigma}_{ii'},\ i,i' = 1,2,\ldots,n\right\}$ and let $\hat{\Omega} = I \otimes \hat{\sigma}_{**}$, then:

$$\hat{\Phi} - \Phi = (y-f)^T \hat{\Omega}^{-1}(y-f) - (y-f)^T \Omega^{-1}(y-f)$$

$$= (y-f)^T(\hat{\Omega}^{-1}-\Omega^{-1})(y-f)$$

and from Lemma 4.2 this converges in probability to zero. Therefore the minimum of $\hat{\Phi}$ converges in probability to the minimum of Φ and also the estimator of θ found by minimizing $\hat{\Phi}$ will converge in probability to the estimator found by minimizing Φ. Moreover under the assumption that the vectors ε_{i*} each have the same multivariate normal distribution, the estimator of θ found by minimizing Φ will correspond to the maximum likelihood estimator. Hence the estimator of θ found by minimizing $\hat{\Phi}$ will converge in probability to the maximum likelihood estimator and thereby is a consistent estimator of θ by using a similar type of discussion as given in Lemma 4.2.

4.4 Modifications to the estimation procedure

We now want to present some modifications to the above generalized least squares estimation procedure that will make it more generally applicable. The situation we consider arises when observations are not available on each equation for each value of the independent variable X. Suppose that we have N_i observations on the i^{th} equation and that all of the N_i are not necessarily equal. Let $N. = \sum_{i=1}^{N} N_i$ and let N equal the total number of different input vectors X_j for all n equations. In general $N_i \leq N$, however $N_i = N$ when the same set of input vectors is used on each of the n regression equations. Let the definitions of y_{*j}, f_{*j}, and ε_{*j} as given earlier apply now as if observations were made for each on all N of the X_j input vectors. For each value of i define the new vector y_{*j}^D which is obtained from the vector y_{*j} by deleting those elements from y_{*j} for which there are no observations. The vector ε_{*j}^D and the matrices f_{*j}^D and Ω^D are defined similarly. Other quantities are constructed from y_{*j}^D, ε_{*j}^D, f_{*j}^D, and Ω^D just as they were from y_{*j}, ε_{*j}, f_{*j}, and Ω Section 4.3, and these quantities are also labeled with the superscript D. Corresponding to equation (4.10), when a different set of input vectors is used on some of the n regression equations we have:

$$\Phi^D = (y^D - f^D)^T \ (\Omega^D)^{-1} \ (y^D - f^D). \tag{4.16}$$

In order to calculate the estimates of the elements $\sigma_{ii'}^D$ of Ω^D, the N_i observations on equation i are used to find the least squares estimator of θ, denoted by $\hat{\theta}^{D(i)}$. Then

$$\hat{\sigma}_{ii}^D = (\hat{\varepsilon}_{i*}^D)^T \ (\hat{\varepsilon}_{i*}^D) \ /N_i. \tag{4.17}$$

To compute $\hat{\sigma}_{ii'}^D$ for $i \neq i'$, let $N_{ii'}$ equal the number of observations on equations i and i′ that have the same X_j input vectors, where we require $N_{ii'} > 0$. The $N_{ii'}$ observations on each of the equations i and i′ are then used to compute separate single equation least squares estimates of θ denoted by $\hat{\theta}^{D(i_{i'})}$ and $\hat{\theta}^{D(i'_i)}$, respectively.

Then

$$\hat{\sigma}^{D}_{ii'} = (\hat{\varepsilon}^{D}_{i_{i'}*})^{T} (\hat{\varepsilon}^{D}_{i'_{i}*}) / N_{ii'} \tag{4.18}$$

where $\hat{\varepsilon}^{D}_{i_{i'}*}$ is computed using $\hat{\theta}^{D(i_{i'})}$, $\hat{\varepsilon}^{D}_{i'_{i}*}$ is computed using $\hat{\theta}^{D(i'_{i})}$, and each is a vector with $N_{ii'}$ elements. Next an iterative procedure for estimating θ is started by computing $\hat{\delta}$ using equation (4.13) after setting ${}_{0}f' = {}_{0}f'^{D}$, $\Omega = \hat{\Omega}^{D}$, and $y = y^{D}$. In order for the limiting properties of these modified estimators to hold we also assume that as $N. \longrightarrow \infty$, $N_{i} \longrightarrow \infty$ and $N_{ii'} \longrightarrow \infty$ for each i and $i' \neq i$.

At this point we will demonstrate the modifications that arise when some of the distributional assumptions are altered. First we will assume that $\Omega = \sigma I$ where $0 < \sigma < \infty$ and I is an $nN \times nN$ identity matrix, i.e. the vector ε is assumed to be made up of independent and identically distributed random variables. To obtain the least squares estimator of θ for this case we minimize:

$$\Phi = \sum_{i=1}^{n} \sum_{j=1}^{N} (y_{ij} - f_{i}(X_{j};\theta))^{2}. \tag{4.19}$$

From an investigation of (4.19) we see that this expression is similar to the expression to be minimized for the single nonlinear regression equation problem. Therefore all of the procedures, properties, and conditions carry over to this case from the single equation case. Obviously under the assumption of normally distributed ε_{ij} for all i and j, the least squares estimator of θ found from minimizing (4.19) is equivalent to the maximum likelihood estimator of θ.

Next let us consider the case where Ω is a diagonal matrix given by:

$$I \otimes \begin{pmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_{nn} \end{pmatrix} = \Omega \tag{4.20}$$

where I is an N × N identity matrix. To obtain the generalized least squares estimator of θ for this case we minimize

$$\Phi = (y-f)^T \, \Omega^{-1} \, (y-f)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{N} \frac{(y_{ij}-f_i(X_j;\theta))^2}{\sigma_{ii}} \quad . \tag{4.21}$$

For the situation where the elements σ_{ii} are assumed to be known, the estimation procedure will be the same as those previously discussed. For the case where the elements σ_{ii} are unknown, then we obtain consistent estimators of σ_{ii} by the same procedure as we discussed earlier using the observations on the i^{th} equation. The estimation of θ is simplified for this case since we do not have to estimate elements of the form $\sigma_{ii'}$ for $i \neq i'$.

For the case when $E(\varepsilon_{ij}\ \varepsilon_{ij'})$ is not independent of j and these covariances are unknown, then instead of allowing only one observation on each equation at each value of X_j we must have more than one observation on each equation at each value of X_j in order to estimate the necessary variances and covariances. Let n_{ij} represent the number of observations made on the i^{th} equation for the input vector X_j, and let all of the n_{ij} be equal, say to ℓ. Then these ℓ observations would be used to obtain an estimator of $E(\varepsilon_{ij}^2)$ in a manner similar to that used in equation (4.17) to obtain an estimator of σ_{ii}. Similarly we find estimators of $E(\varepsilon_{ij}\ \varepsilon_{ij'})$, $E(\varepsilon_{ij}\ \varepsilon_{i'j})$, and $E(\varepsilon_{ij}\ \varepsilon_{i'j'})$ where $i \neq i'$ and $j \neq j'$. With consistent estimators now provided for the covariance matrix, we proceed to estimate the vector θ by the same method we described earlier.

Another example, similar to one discussed earlier, except that multiple observations at each value of X_j on each equation would not necessarily be required, would be the case when $E(\varepsilon_{ij}^2) = f_i(X_j;\theta)$, and $E(\varepsilon_{ij}\varepsilon_{i'j}) = E(\varepsilon_{ij}\varepsilon_{ij'}) = E(\varepsilon_{ij}\varepsilon_{i'j'}) = 0$ for $i,i' = 1,2,\ldots,n$, $j,j' = 1,2,\ldots,N$, $i \neq i'$, and $j \neq j'$. This is a more general example of equations (4.20) and (4.21), and it can be shown for this particular case that Φ reduces to

$$\Phi = \sum_{i=1}^{n} \sum_{j=1}^{N} \frac{(y_{ij}-f_i(X_j;\theta))^2}{f_i(X_j;\theta)} \quad . \tag{4.22}$$

This situation might arise when our random variables where related to a Poisson distributed random variable. Although some of the above cases become much more involved than the originally discussed case, these modifications do show that it is possible to apply this generalized least squares estimation technique to a variety of classes of distributions for the vector ε.

V. A GENERALIZED PARTIAL TOTALS ESTIMATION PROCEDURE

5.1 Introduction

As we indicated in Chapters 1 and 3 there are numerous examples of experimental situations in the fields of biology, chemistry, and physics that yield data which are easily described by equations of the form (1.1). Cornell [1956, 1962] develops an estimation procedure based on the concept of partial totals that may be used in estimating the parameters in equation (1.1) for the case $n = 1$. In Section 5.2 we will outline this partial totals estimation procedure for the sake of completeness, since the results presented there will be used in succeeding sections.

In Chapter 3 we demonstrated for the $(n+1)$-compartment mammillary and catenary models that the n independent equations describing the experimental situation are each a linear combination of the same n exponential terms, that is, for equation (1.1) that $m = n$ in this situation. Section 5.3 will be devoted to the development and description of a generalized partial totals estimation procedure for the regression model of the type given by (1.1) for the case when $n = m$ and the values of the independent variable x_j are equally spaced. In Section 5.5 certain alternatives will be suggested for the cases when some of the observations are not taken at equally spaced values of x_j. In Section 5.4 some theorems will be presented concerning some of the properties of the estimators found by this generalized estimation procedure.

5.2 Single equation partial totals estimation

The results in this section are contained in the work by Cornell [1956, 1962]; however, we repeat them here since some of the results will be needed for the generalized procedure to be presented in the next section. In this section we will again suppress the subscript i. Let Y_j be an observable random variable where

$$Y_j = \sum_{k=1}^{m} \alpha_k e^{-\lambda_k x_j} + \varepsilon_j = E(Y_j) + \varepsilon_j, \tag{5.1}$$

for j = 0,1,2,...,2Mm-1 where M is a positive integer. We will assume that $x_j = hj$ for all j where h is a positive constant, so $x_o = 0$ and $x_{j+1} - x_j = h$. Next we form the following partial totals:

$$\Sigma_q = \sum_{j=(q-1)M}^{qM-1} E(Y_j) = \sum_{j=(q-1)M}^{qM-1} \sum_{k=1}^{m} \alpha_k e^{-\lambda_k hj}, \tag{5.2a}$$

for q = 1,2,...,2m. Since $\sum_{j=(q-1)M}^{qM-1} e^{-\lambda_k hj}$ is a geometric series, we may now write (5.2a) as:

$$\Sigma_q = \sum_{k=1}^{m} \alpha_k e^{-\lambda_k h(q-1)M} \frac{(1-e^{-\lambda_k hM})}{(1-e^{-\lambda_k h})} . \tag{5.2b}$$

Let $\Lambda_1, \Lambda_2, \ldots, \Lambda_m$ represent the m elementary symmetric functions of $e^{-\lambda_1 hM}, e^{-\lambda_2 hM}, \ldots, e^{-\lambda_m hM}$, i.e. Λ_r is the sum of all possible distinct products of the $e^{-\lambda_k hM}$ taken r at a time. If we let $\Lambda_o = 1$, then we can show that the following set of equations is satisfied:

$$\sum_{r=1}^{m+1} (-1)^{2m+1-r} \Lambda_{m+1-r} \Sigma_{q+r} = 0 \tag{5.3}$$

for $q = 0,1,\ldots,m-1$.

We note that (5.3) is a set of m equations which are linear in the m unknowns $\Lambda_1, \Lambda_2, \ldots, \Lambda_m$. Therefore we next set the quantities Σ_q equal to the corresponding observed partial totals $S_q = \sum_{j=(q-1)M}^{qM-1} y_j$ for $q = 1,2,\ldots,2m$ where y_j represents the observed value of the r.v. Y_j, and we solve the resulting set of linear equations corresponding to (5.3). The solutions to this set of linear equations give us estimators L_r of the symmetric functions Λ_r, $r = 1,2,\ldots,m$. Since the L_r estimate the symmetric functions of $e^{-\lambda_k hM}$, estimators of the quantities $e^{-\lambda_k hM}$ may be found by obtaining the m roots of the polynomial equation:

$$w^m - L_1 w^{m-1} + L_2 w^{m-2} - \ldots + (-1)^m L_m = 0. \tag{5.4}$$

Let the roots of this equation be denoted by $w_1, w_2, \ldots, w_m$. Then the estimators of the λ_k are given by $\hat{\lambda}_k = -\frac{1}{Mh} \ln w_k$ for $k = 1,2,\ldots,m$.

To obtain estimators of the parameters $\alpha_1,\ldots,\alpha_m$ the following subset of equation (5.2b) is solved by elementary methods for $\hat{\alpha}_1,\ldots,\hat{\alpha}_m$:

$$S_q = \sum_{k=1}^{m} \hat{\alpha}_k e^{-\hat{\lambda}_k h(q-1)M} \frac{1-e^{-\hat{\lambda}_k hM}}{1-e^{-\hat{\lambda}_k h}}, \quad q = 1,2,\ldots,m. \tag{5.5}$$

Cornell [1956, 1962] also develops the partial totals estimation procedure for the following regression model:

$$Y_j = \alpha_o + \sum_{k=1}^{m} \alpha_k e^{-\lambda_k hj} + \varepsilon_j \tag{5.6}$$

for $j = 0,1,\ldots,(2m+1)M - 1$. For this case we form the differences

$$\Sigma_q - \Sigma_{q+1} = \sum_{k=1}^{m} \alpha_k e^{-\lambda_k h(q-1)M} (1-e^{-\lambda_k hM})^2/(1-e^{-\lambda_k h}) \tag{5.7}$$

for $q = 1,2,\ldots,2m+1$. After substituting $S_q - S_{q+1}$ for $\Sigma_q - \Sigma_{q+1}$, the solution for the estimators L_r of Λ_r is the same in terms of the differences $S_q - S_{q+1}$ as that given by the solutions of (5.3) in terms of the S_q substituted in the place of Σ_q. The estimators of $\alpha_1,\alpha_2,\ldots,\alpha_m$ are found by substituting S_q and $\hat{\lambda}_k$ in place of Σ_q and λ_k respectively in the first m equations of (5.7). Finally, an estimator of α_o may be found by substituting S_1, $\hat{\lambda}_k$, and $\hat{\alpha}_k$ for Σ_1, λ_k, and α_k, respectively, in the following equation:

$$\Sigma_1 = M\alpha_o + \sum_{k=1}^{m} \alpha_k \frac{(1-e^{-\lambda_k hM})}{(1-e^{-\lambda_k h})}. \tag{5.8}$$

Moreover, Cornell also proves that the estimators of the parameters in (5.1) and (5.6) found by this partial totals approach are consistent estimators under the following assumptions:

1) The random variables ε_j are independent for all values of j.

2) The random variables ε_j are identically distributed for all values of j in the same group or partial total.

3) The domain of the independent variable $x_j = hj$ remains constant for each group or partial total as $M \to \infty$, i.e. as $M \to \infty$ we must have $h \to 0$ but Mh remaining constant. Finally, the asymptotic normality of the distributions of these partial totals estimators is demonstrated by Cornell.

5.3 Description and development of the generalized partial totals estimation procedure

Although this estimation procedure was motivated by the consideration of the regression equations that arise when we are concerned with tracer experiments, this section will present the estimation technique for two more general cases; then by a reference to Theorems 3.1 and 3.2 of Chapter 3, we see that these models relate to the tracer experiment problem of interest. The two more general cases may be described as follows:

Case I: The regression model is given by:

$$Y_{ij} = \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k x_j} + \varepsilon_{ij}, \tag{5.9}$$

for $i = 1,2,\ldots,n$ and $j = 0,1,\ldots,(n+1)M-1$, where the observable random variable Y_{ij} takes on values denoted by y_{ij}.

Case II: The regression model is given by:

$$Y_{ij} = \alpha_{io} + \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k x_j} + \varepsilon_{ij}, \tag{5.10}$$

for $i = 1,2,\ldots,n$ and $j = 0,1,\ldots,(n+2)M-1$, where the observable random variable Y_{ij} takes on values denoted by y_{ij}.

In the above two cases we are assuming that n and M are positive integers, the coefficients α_{ik} are real numbers, and the exponents λ_k are distinct positive real numbers. Since we will want to take our observations at equally spaced values of the independent variable x_j, we will assume that $x_j = hj$ where h is a positive constant.

The estimation of the exponential parameters will involve the application of a partial totals approach similar to that discussed by Cornell [1956, 1962]. First we will consider the estimation of the exponential parameters for the regression model given by Case I. The first step will involve the grouping of the observations from each equation into (n+1) groups each containing M observations, and then the formulation of the following partial totals:

$$\sum_{j=(q-1)M}^{qM-1} Y_{ij} = \sum_{j=(q-1)M}^{qM-1} \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k hj} + \sum_{j=(q-1)M}^{qM-1} \varepsilon_{ij}$$

$$= \Sigma_{iq} + \sum_{j=(q-1)M}^{qM-1} \varepsilon_{ij} \tag{5.11}$$

for $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$ where $\Sigma_{iq} = \sum_{j=(q-1)M}^{qM-1} E(Y_{ij})$.

Now for each value of i we can use the same steps that were used to derive equation (5.3) to show that the following equation is satisfied:

$$\Lambda_n\Sigma_{i1} - \Lambda_{n-1}\Sigma_{i2} + \Lambda_{n-2}\Sigma_{i3} - \ldots + (-1)^n\Lambda_o\Sigma_{i,n+1} = 0 \qquad (5.12)$$

for $i = 1,2,\ldots,n$ where Λ_r, $r = 1,2,\ldots,n$, are the elementary symmetric functions of $e^{-\lambda_k hM}$, i.e. they equal the sum of all possible products of the terms $e^{-\lambda_k hM}$ taken r at a time. In addition, we define $\Lambda_o = 1$. Therefore, since $i = 1,2,\ldots,n$, in equation (5.12) we have n equations in the n unknowns Λ_r. Hence by substituting $S_{iq} = \sum_{j=(q-1)M}^{qM-1} y_{ij}$ for Σ_{iq} we may easily solve for estimators of Λ_r which will be denoted by L_r. From these estimators of the elementary symmetric functions we may now obtain estimators of $e^{-\lambda_k hM}$ for $k = 1,2,\ldots,n$. Using the same properties of elementary symmetric functions that we used to derive equation (5.4), the estimators of $e^{-\lambda_k hM}$, $k = 1,2,\ldots,n$, are obtained by finding the n roots of the polynomial equation:

$$w^n - L_1w^{n-1} + L_2w^{n-2} - \ldots + (-1)^n L_n = 0 . \qquad (5.13)$$

Let the roots of (5.13) be denoted by $w_1, w_2, \ldots, w_n$. Then the estimators of λ_k are given by $\hat{\lambda}_k = \left(\frac{-1}{hM}\right) \ln w_k, k = 1,2,\ldots,n$.

For Case II given by equation (5.10) we group the observations into (n+2) groups each containing M observations, and then form the following partial totals: $S'_{iq} = S_{iq} - S_{i,q+1}$ for $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$, where the partial totals $S_{iq} = \sum_{j=(q-1)M}^{qM-1} y_{ij}$. For each S'_{iq} there is a corresponding $\Sigma'_{iq} = \Sigma_{iq} - \Sigma_{i,q+1}$, and we can show that the following equation is satisfied by the Σ'_{iq}:

$$\Lambda_n \Sigma'_{i1} - \Lambda_{n-1}\Sigma'_{i2} + \Lambda_{n-2}\Sigma'_{i3} - \ldots + (-1)^n \Lambda_o \Sigma'_{i,n+1} = 0, \qquad (5.14)$$

$i = 1,2,\ldots,n$, i.e., this set of equations is the same as (5.12) except that Σ'_{iq}'s have been substituted for Σ_{iq}'s. The Λ_r, $r = 0,1,2,\ldots,n$, are the same as those defined earlier. We now proceed as in Case I to obtain estimators of the exponential parameters using the S'_{iq}'s instead of the S_{iq}'s, $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$.

At this point we want to estimate the coefficients in the set of equations (5.9). To obtain these estimators we will substitute the estimators of λ_k, found by the partial totals procedure described above and denoted by $\hat{\lambda}_k$, into our set of n independent equations and use a least squares procedure to estimate the unknown linear coefficients. We will proceed as if there did not exist any functional relationships between the exponential parameters and the linear coefficients. If the exact relation is known between the exponential parameters and the linear coefficients and if the linear coefficients are completely specified by the exponential parameters, then the values of the estimators of the exponential parameters can be substituted into these relations giving us estimates of the linear coefficients. If the exact relation is known between the exponential parameters and the linear coefficients but the linear coefficients are not completely specified by the exponential parameters, then the values of the estimators of the exponential parameters can be substituted into these relations giving us a set of regression equations for the remaining parameters. If the system of equations is linear then we can proceed in the same manner as described below, and if the system of regression equations is nonlinear then we can use an iterative technique such as that described in Chapter 4.

In order to reduce the amount of space needed to write the necessary equations, we will use matrix notation. Therefore we will need to define the following vectors and matrices for $i = 1,2,\ldots,n$:

$y_{**} = (y_{1*}^T, y_{2*}^T, \ldots, y_{n*}^T)^T$, where $y_{i*}^T = (y_{io}, y_{i1}, \ldots, y_{i,(n+1)M-1})$;

$\alpha_{**} = (\alpha_{1*}^T, \alpha_{2*}^T, \ldots, \alpha_{n*}^T)^T$, where $\alpha_{i*}^T = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$;

$\varepsilon_{**} = (\varepsilon_{1*}^T, \varepsilon_{2*}^T, \ldots, \varepsilon_{n*}^T)^T$, where $\varepsilon_{i*}^T = (\varepsilon_{io}, \varepsilon_{i1}, \ldots, \varepsilon_{i,(n+1)M-1})$;

and

$$Z = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ e^{-\lambda_1 h} & e^{-\lambda_2 h} & \cdots & e^{-\lambda_n h} \\ e^{-2\lambda_1 h} & e^{-2\lambda_2 h} & \cdots & e^{-2\lambda_n h} \\ \vdots & \vdots & & \vdots \\ e^{-[(n+1)M-1]\lambda_1 h} & e^{-[(n+1)M-1]\lambda_2 h} & \cdots & e^{-[(n+1)M-1]\lambda_n h} \end{pmatrix} .$$

Using the above definitions we may now write the complete set of equations given by (5.9) as follows:

$$y_{**} = D_Z \alpha_{**} + \varepsilon_{**} \tag{5.15}$$

where D_Z is an $[n(n+1)M] \times n^2$ matrix with Z matrices along its diagonal, i.e.,

$$D_Z = \begin{pmatrix} Z & 0 & \cdots & 0 \\ 0 & Z & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & Z \end{pmatrix}$$

where the 0's in the above matrix are matrices of the appropriate dimensions with only zeros as elements.

Let Ω be the covariance matrix of the vector ε_{**} which we discussed in Section 3.2. If we knew the elements of Ω and the true values of the exponential parameters, then the usual weighted least squares estimator of the vector α_{**} would be taken as:

$$\hat{\alpha}_{**} = (D_Z^T \Omega^{-1} D_Z)^{-1} (D_Z^T \Omega^{-1} y_{**}). \tag{5.16}$$

Since, in reality, we do not know the elements of the matrix Ω or the true values of the exponential parameters, we will substitute our partial totals estimators of the exponential parameters into the i^{th} regression equation and obtain the usual least squares estimators of the linear parameters in the vector α_{i*} by

$$\hat{\alpha}_{i*} = (\hat{Z}^T\hat{Z})^{-1}\hat{Z}^T y_{i*},$$

where $\hat{Z}$ is the matrix Z with the partial totals estimators substituted in the place of the unknown parameter values. Doing this for each value of $i = 1,2,\ldots,n$, we find the estimators of ε_{i*} by

$$\hat{\varepsilon}_{i*} = y_{i*} - \hat{Z}\,\hat{\alpha}_{i*}.$$

From Section 5.4 we note that we will take $\Omega = \Sigma \otimes I$ where $\Sigma = \{\sigma_{ii'};i,i' = 1,2,\ldots,n\}$ is a positive definite matrix. The estimators of $\sigma_{ii'}$ are given by

$$\hat{\sigma}_{ii'} = \hat{\varepsilon}_{i*}^T\hat{\varepsilon}_{i'*}/(n+1)M, \tag{5.17.1}$$

for $i,i' = 1,2,\ldots,n$. Our estimator of Ω is found by substituting these estimators into Σ, giving us $\hat{\Sigma}$, and then taking $\hat{\Omega} = \hat{\Sigma} \otimes I$. Using $\hat{\Omega}$ we will have the following expression for the estimator of the vector $\hat{\alpha}_{**}$:

$$\hat{\hat{\alpha}}_{**} = (D_{\hat{Z}}^T\hat{\Omega}^{-1}D_{\hat{Z}})^{-1}(D_{\hat{Z}}^T\hat{\Omega}^{-1}y_{**}). \tag{5.17}$$

For the regression model specified by equation (5.10), which we have designated as Case II, we will need to define the following new set of vectors and matrices:

$$y'_{**} = (y'^{T}_{1*}, y'^{T}_{2*}, \ldots, y'^{T}_{n*})^{T}, \text{ where } y'^{T}_{i*} = (y_{io}, y_{i1}, \ldots, y_{i,(n+2)M-1});$$

$$\alpha'_{**} = (\alpha'^{T}_{1*}, \alpha'^{T}_{2*}, \ldots, \alpha'^{T}_{n*})^{T}, \text{ where } \alpha'^{T}_{i*} = (\alpha_{io}, \alpha_{i1}, \ldots, \alpha_{in});$$

$$\varepsilon'_{**} = (\varepsilon'^{T}_{1*}, \varepsilon'^{T}_{2*}, \ldots, \varepsilon'^{T}_{n*})^{T}, \text{ where } \varepsilon'^{T}_{i*} = (\varepsilon_{io}, \varepsilon_{i1}, \ldots, \varepsilon_{i,(n+2)M+1});$$

and

$$Z' = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & e^{-\lambda_1} & e^{-\lambda_2} & \cdots & e^{-\lambda_n} \\ 1 & e^{-2\lambda_1} & e^{-2\lambda_2} & \cdots & e^{-2\lambda_n} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & e^{-[(n+2)M-1]h\lambda_1} & e^{-[(n+2)M-1]h\lambda_2} & \cdots & e^{-[(n+2)M-1]h\lambda_n} \end{pmatrix} .$$

From the above definitions we can now write the complete set of equations given by (5.10) as follows:

$$y'_{**} = D_{Z'}\alpha'_{**} + \varepsilon'_{**} \tag{5.18}$$

where $D_{Z'}$ is an $[n(n+2)M \times n(n+1)]$ matrix with Z' matrices along its diagonal, i.e.

$$D_{Z'} = \begin{pmatrix} Z' & 0 & \cdots & 0 \\ 0 & Z' & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & Z' \end{pmatrix}$$

where 0's have the same meaning as given for Case I.

Using a similar type of reasoning as was applied in Case I, we find an estimator of the vector α'_{**} of linear parameters to be given by the following expression:

$$\hat{\hat{\alpha}}'_{**} = (D^T_{\hat{Z}'} \hat{\Omega}'^{-1} D_{\hat{Z}'})^{-1} (D^T_{\hat{Z}'} \hat{\Omega}'^{-1} y'_{**}), \tag{5.19}$$

where the ^'s have the same meaning here as they had in Case I.

5.4 Some properties of the generalized partial totals estimators

During the development of the generalized estimation procedure presented in Section 5.3, the only assumption that we used concerning the random variables ε_{ij} was the assumption that $E(\varepsilon_{ij}) = 0$ for all i and j. However, before we can investigate some of the properties of these estimators we must make some more specific assumptions about the random variables ε_{ij}. These assumptions may be stated as follows:

1) For each value of i and j we have $E(\varepsilon_{ij}) = 0$ and $E(\varepsilon^2_{ij}) = \sigma_{ii}$ where $0 < \sigma_{ii} < \infty$.

2) For each value of i, i', j, and j' with $i \neq i'$ and $j \neq j'$ we have $E(\varepsilon_{ij}\varepsilon_{ij'}) = E(\varepsilon_{ij}\varepsilon_{i'j'}) = 0$ and $E(\varepsilon_{ij}\varepsilon_{i'j}) = \sigma_{ii'}$ where $-\infty < \sigma_{ii'} < \infty$.

At this point we will prove the Theorem 5.1 given below, in order to demonstrate the consistency of the estimators of the exponential parameters for Case I. After proving this theorem we will indicate the minor changes to be made in the proof in order to demonstrate the consistency of the estimators for the exponential parameters for Case II.

Theorem 5.1: Let $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n$ be the generalized partial totals estimators of the parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$ in the regression model given by (5.9), which we have designated as Case I. These estimators of the exponential parameters are consistent estimators under the following assumptions:

1) For each value of $i = 1,2,\ldots,n$, the random variables ε_{ij}, $j = 0,1,\ldots,(n+1)M-1$, are uncorrelated with $E(\varepsilon_{ij}) = 0$.

2) For each value of i and q the random variables ε_{ij} associated with the corresponding observations y_{ij} in S_{iq} as given in Section 5.3 have constant variance.

3) For each value of i and q the domain of the independent variable is of constant length J for S_{iq} where $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$.

4) For the linear coefficients α_{ik}, let α be the $n \times n$ matrix of these coefficients, where α_{ik} is the element in the i^{th} row and k^{th} column for $i, k = 1,2,\ldots,n$, and assume that the determinant of α, $|\alpha|$, is unequal to zero.

Proof: From the substitution of the S_{iq} for the Σ_{iq} in the set of equations displayed in (5.12), we note that the estimators L_r of the Λ_r, $r = 1,2,\ldots,n$, are found by Cramer's rule as the ratio of the following two determinants:

$$L_r = |P_r|/|P| \ , \quad r = 1,2,\ldots,n, \tag{5.20a}$$

where P is an $n \times n$ matrix whose $(i,q)^{th}$ element is $(-1)^{q-1}S_{iq}$ for i, q = 1,2,...,n. The $n \times n$ matrix P_r is the same as P except that the elements in the $(n-r+1)^{th}$ column are replaced by the elements $(-1)^{n-1}S_{i,n+1}$. Now since each S_{iq} is the sum of M observations, we may replace each S_{iq} by $\overline{S}_{iq}$, where $\overline{S}_{iq} = S_{iq}/M$, and still have the same estimators for L_r. Therefore let

$$L_r = |R_r|/|R| \ , \quad r = 1,2,\ldots,n, \tag{5.20b}$$

where R_r and R are respectively the same as P_r and P except with the $\overline{S}_{iq}$'s substituted for the S_{iq}'s. Next let us write $\overline{S}_{iq}$ as follows:

$$\overline{S}_{iq} = \frac{1}{M} \sum_{j=(q-1)M}^{qM-1} y_{ij} = \frac{1}{M} \sum_{j=(q-1)M}^{qM-1} \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k hj} + \frac{1}{M} \sum_{j=(q-1)M}^{qM-1} \varepsilon_{ij}. \tag{5.21a}$$

We will allow $M \to \infty$. However, as stated in the third assumption of our theorem, we will specify that the M observations included in the q^{th} partial total for a particular equation are made for the values of j given by

$$(q-1)J, (q-1)J + {}^{J}/M, (q-1)J + {}^{2J}/M, \ldots, {}^{J}/M(qM-1).$$

Therefore, with this specification, equation (5.21a) may be written as

$$\overline{S}_{iq} = \frac{1}{M} \sum_{j=(q-1)M}^{qM-1} \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k {}^{Jj}/M} + \frac{1}{M} \sum_{j=(q-1)M}^{qM-1} \varepsilon_{ij}. \tag{5.21b}$$

From the first and second assumptions of our theorem, the last term in (5.21b) is the mean of M uncorrelated random variables each with the same variance. By an application of the Tchebycheff theorem given in Cramér ([1946], page 253), this term will converge in probability to $E(\varepsilon_{ij}) = 0$.

$$\lim_{M\to\infty} \frac{1}{J} \sum_{j=(q-1)M}^{qM-1} \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k Jj/M}\, J/M = \frac{1}{J} \sum_{k=1}^{n} \int_{(q-1)J}^{qJ} \alpha_{ik} e^{-\lambda_k x} dx$$

$$= \frac{1}{J} \sum_{k=1}^{n} \frac{\alpha_{ik}}{\lambda_k} e^{-\lambda_k (q-1)J} (1-e^{-\lambda_k J}). \tag{5.22}$$

From Slutsky's theorem (see Cramér [1946], page 255) we now have that $\bar{S}_{iq}$ converges in probability to the above constant, which will be denoted by Ψ_{iq}, i.e.

$$\bar{S}_{iq} \xrightarrow[p]{} \Psi_{iq} = \frac{1}{J} \sum_{k=1}^{n} \frac{\alpha_{ik}}{\lambda_k} e^{-\lambda_k (q-1)J} (1-e^{-\lambda_k J}) \tag{5.23a}$$

as $M \to \infty$ where $\xrightarrow[p]{}$ denotes convergence in probability.

Now in order to complete the demonstration of $\hat{\lambda}_k \xrightarrow[p]{} \lambda_k$ as $M\to\infty$ for $k = 1,2,\ldots,n$, we will need to go through an argument similar to the one given by Cornell [1956] for one regression equation. From equation (5.20b) we note that the estimators L_r are merely ratios of sums of products of the $\bar{S}_{iq}$. Therefore $|R_r|$ and $|R|$ are continuous functions of the $\bar{S}_{iq}$, and L_r will be continuous at the point Ψ_{**}, where Ψ_{**} is defined as that point where $\bar{S}_{iq} = \Psi_{iq}$ for all i and q, provided $|R| \neq 0$ at Ψ_{**}. We may write Ψ_{iq} as follows:

$$\Psi_{iq} = \sum_{k=1}^{n} u_{ik} e^{-\lambda_k (q-1)J} \tag{5.23b}$$

where $u_{ik} = \dfrac{\alpha_{ik}}{J\lambda_k}(1-e^{-\lambda_k J})$ for i, k = 1,2,...,n and q = 1,2,...,n+1.

For all i and k we note that $u_{ik} \neq 0$. Now at the point Ψ_{**} the $(i,q)^{th}$ element of R is given by (5.23b) and therefore at this point R may be written as the following product:

$$UW = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} 1 & e^{-\lambda_1 J} & \cdots & e^{-\lambda_1 (n-1)J} \\ 1 & e^{-\lambda_2 J} & \cdots & e^{-\lambda_2 (n-1)J} \\ \vdots & \vdots & & \vdots \\ 1 & e^{-\lambda_n J} & \cdots & e^{-\lambda_n (n-1)J} \end{pmatrix}. \tag{5.24}$$

The matrix W is a Vandermonde matrix (see Hohn [1964], page 70) and therefore $|W| = \prod_{k>k'} (e^{-\lambda_k J} - e^{-\lambda_{k'} J})$. Moreover, it can be shown that $|U| = |\alpha| \prod_{k=1}^{n} \dfrac{(1-e^{-\lambda_k J})}{\lambda_k J}$. From the fourth assumption of our theorem we note that $|R|$ at the point Ψ_{**} is unequal to zero, which implies that L_r, r = 1,2,...,n, is continuous in a neighborhood of Ψ_{**}. Now since $w_1, w_2, \ldots, w_n$ are the n roots of an n^{th} degree polynomial with coefficients L_r, then $w_1, w_2, \ldots, w_n$ will be continuous in a neighborhood of Ψ_{**} since the roots of a polynomial are continuous functions of the coefficients. Now the estimators $\hat{\lambda}_k$ are continuous functions of the roots w_k, k = 1,2,...,n. Therefore we have that the estimators of

λ_k, $k = 1,2,\ldots,n$, are continuous in a neighborhood of Ψ_{**}.

Now that we have demonstrated that the estimators $\hat{\lambda}_k$, $k = 1,2,\ldots,n$, are continuous in a neighborhood of Ψ_{**}, we may apply a result proved by Slutsky (see Sverdrup [1952], page 6) to conclude that if $\hat{\lambda}_k\big|_{\overline{S}_{iq}=\Psi_{iq}} = \lambda_k$, then $\hat{\lambda}_k$ converges in probability to λ_k with $\overline{S}_{iq} = \Psi_{iq}$ for all i and q. So in order to complete the demonstration that the $\hat{\lambda}_k$, $k = 1,2,\ldots,n$, are consistent we must show that $\hat{\lambda}_k = \lambda_k$ at the point Ψ_{**} for $k = 1,2,\ldots,n$. Let $\zeta_{iq} = E(\overline{S}_{iq})$. Then from equations (5.21b) and (5.22) we note that $\zeta_{iq} \to \Psi_{iq}$ as $M\to\infty$. From equation (5.12) we can see that $L_r = \Lambda_r$, $r = 1,2,\ldots,n$, at the point $\overline{S}_{iq} = \zeta_{iq}$ for all i and q. This implies that the roots of the polynomial equation

$$w^n - \Lambda_1 w^{n-1} + \Lambda_2 w^{n-2} - \ldots + (-1)^n \Lambda_n = 0 \tag{5.25}$$

would be $e^{-\lambda_k hM} = e^{-\lambda_k J}$, $k = 1,2,\ldots,n$. Therefore we have that $\hat{\lambda}_k = \lambda_k$ at the point $\overline{S}_{iq} = \zeta_{iq}$ for all i and q. Using the above conclusions we have

$$\hat{\lambda}_k\Big|_{\overline{S}_{iq}=\Psi_{iq}} = \lim_{M\to\infty} \hat{\lambda}_k\Big|_{\overline{S}_{iq}=\zeta_{iq}} = \lambda_k, \tag{5.26}$$

where $\overline{S}_{iq} = \Psi_{iq}$ and $\overline{S}_{iq} = \zeta_{iq}$ are to hold for all i and q. Hence $\hat{\lambda}_k$ converges in probability to λ_k for $k = 1,2,\ldots,n$, and $\hat{\lambda}_k$ is by definition a consistent estimator of λ_k. This completes the proof of Theorem 5.1.

For the regression model designated as Case II by equation (5.10) we have the following theorem:

Theorem 5.2: Let $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n$ be the generalized partial totals estimators of the parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$ in the regression model given by (5.10), which we have designated as Case II. These estimators of the exponential parameters are consistent estimators under the following assumptions:

1) For each value of $i = 1,2,\ldots,n$, the random variables ε_{ij}, $j = 0,1,2,\ldots,(n+2)M-1$, are uncorrelated with $E(\varepsilon_{ij}) = 0$.

2) For each value of i and q the random variables ε_{ij} associated with the corresponding observations y_{ij} in S_{iq}, as given in Section 5.3, have constant variance.

3) For each value of i and q, the domain of the independent variable is of constant length J for S_{iq} where $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+2$.

4) For the linear coefficients α_{ik}, i, $k = 1,2,\ldots,n$, let α be the $n \times n$ matrix of these coefficients, where α_{ik} is the element in the i^{th} row and k^{th} column, and assume that the determinant of α, $|\alpha|$, is unequal to zero.

Since the proof of this theorem will be similar to the proof of Theorem 5.2, we will merely indicate those points where changes need to be made. In order to show that the estimators of the exponential parameters for Case II are also consistent, we will need the following expression:

$$\frac{1}{M} S'_{iq} = \bar{S}'_{iq} \xrightarrow{p} \frac{1}{J} \sum_{k=1}^{n} \frac{\alpha_{ik}}{\lambda_k} e^{-\lambda_k (q-1) J} (1-e^{-\lambda_k J})^2 = \Psi'_{iq} \tag{5.27}$$

for $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$ as $M\to\infty$. With this expression we will need to show that $|R'|$, i.e. the determinant of R with its elements $\bar{S}_{iq}$ replaced by $\bar{S}'_{iq}$, is unequal to zero at the point Ψ'_{**}. At the point Ψ'_{**} the determinant of R' may be written as the product $|U'||W|$ where the matrix W is the same as was defined in the proof of Theorem 5.1 and

$$|U'| = |\alpha| \prod_{k=1}^{n} \frac{(1-e^{-\lambda_k J})^2}{\lambda_k J} . \tag{5.28}$$

Therefore using the assumptions of our theorem and going through the same continuity argument that was used for Case I, we conclude that the estimators of the exponential parameters for Case II are also consistent.

Before we demonstrate the consistency of the estimators of the linear parameters for Cases I and II, we will need to prove the following lemma.

Lemma 5.3: Let $\hat{a} = (\hat{a}_1,\ldots,\hat{a}_n)^T$ be a vector of consistent estimators of the elements of the vector $a = (a_1,a_2,\ldots,a_n)^T$. Let $b = G(a)$ be a jointly continuous function of the elements of the vector a, such that $\hat{b} - b \xrightarrow{p} 0$. Then $\hat{\hat{b}} - b \xrightarrow{p} 0$ and $\hat{\hat{b}}$ is a consistent estimator of b, where $\hat{\hat{b}}$ is equal to the function $G(\hat{a})$, i.e. G(a) with $\hat{a}$ substituted in the place of a.

Proof: Since $\hat{a}$ is a vector of consistent estimators of a, then we know that $\hat{a}_k - a_k \overrightarrow{p} 0$ for $k = 1,2,\ldots,n$. Since $G(a)$ is jointly continuous in the elements of a, we may apply the results proved by Pratt([1959],pages 551,552) to conclude that $G(\hat{a}) - G(a) \overrightarrow{p} 0$. From our hypothesis we also have that $\hat{b} - b = G(a) - b \overrightarrow{p} 0$, therefore

$$\hat{\hat{b}} - b = G(\hat{a}) - G(a) + G(a) - b \overrightarrow{p} 0$$

since both portions of the sum tend in probability to zero. Hence $\hat{\hat{b}}$ is a consistent estimator of b.

If the exponential parameters of our regression model are known and the random variables ε_{ij} are assumed to be normally distributed, then the estimators of the linear parameters given by equation (5.16) and the corresponding equation for Case II are maximum likelihood estimators of these parameters. The conditions given by Cramér ([1946], page 500) are satisfied by the normal density function and therefore the asymptotic properties of maximum likelihood estimators demonstrated by Cramér hold for this case. In particular, the maximum likelihood estimators converge in probability to the true values of the parameters when the conditions are satisfied. Hence by the use of Lemma 5.3 along with Theorems 5.1 and 5.2 we may prove the following theorem:

Theorem 5.4: Let the following assumptions be satisfied:

1) The assumptions of Theorem 5.1 (or 5.2) are satisfied.
2) The random variables ε_{ij} are normally distributed.
3) The elements of the vector $\hat{\alpha}_{**}$ (or $\hat{\hat{\alpha}}_{**}$) in equation (5.16)

(or (5.19)) are continuous functions.

Then the estimators of the linear parameters given by equation (5.16) (or(5.19)) are consistent estimators of the linear parameters in our regression model.

Now that we have established the consistency of our generalized partial totals estimators, we will investigate the limiting distribution of the exponential estimators and then we will derive an expression for the asymptotic efficiency of these estimators. Before presenting the detailed discussion of this distribution theory, we will introduce some notation that will be used in the development given below. When we have two arbitrary vectors $a = (a_1, a_2, \ldots, a_n)^T$ and $b = (b_1, b_2, \ldots, b_n)^T$, then by $|a|$ we mean the vector $(|a_1|, |a_2|, \ldots, |a_n|)^T$, by $a < b$ we mean $a_i < b_i$ for all i, by $\lim_{N\to\infty} a$ we mean the vector $\left(\lim_{N\to\infty} a_1, \lim_{N\to\infty} a_2, \ldots, \lim_{N\to\infty} a_n\right)^T$. Since the following results will hold for both Cases I and II with obvious modifications, the details of the derivations will be presented only for Case I with merely the conclusions for Case II being shown.

Theorem 5.5: Let $\hat{\lambda}$ represent the $n \times 1$ vector of generalized partial totals estimators for the exponential parameters given by equation (5.9), where each $\hat{\lambda}_k$, $k = 1,2,\ldots,n$, is a function of the $\bar{S}_{iq}$, $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$. If each $\hat{\lambda}_k$ possesses continuous second order derivatives of every kind with respect to the $\bar{S}_{iq}$ in the neighborhood $|\bar{S}_{iq} - \Psi_{iq}| \leq \delta$ for $\delta > 0$

where $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$, and, in addition, if the distributional assumptions concerning the random variables ε_{ij} stated at the beginning of Section 5.4 and Theorem 5.1 are satisfied, then the limiting distribution of $\sqrt{M}(\hat{\lambda}-\lambda_*)$, where λ_* represents the vector $\hat{\lambda}$ with the $\bar{S}_{iq}$ replaced by the Ψ_{iq} for all i and q in each of the elements $\hat{\lambda}_k$, is a multivariate normal distribution with mean vector given by the zero vector and covariance matrix given by: $F\Omega F^T$ where $\Omega = ME(\bar{\varepsilon}_{*}.\bar{\varepsilon}_{*}^T.)$,

$$\bar{\varepsilon}_{*}. = \frac{1}{M}\sum_{j=o}^{M-1} \varepsilon_{*j},\ \varepsilon_{*j} = (\varepsilon_{1j},\varepsilon_{1,j+M},\ldots,\varepsilon_{1,j+nM},\varepsilon_{2j},\ldots,$$

$$\varepsilon_{2,j+nM},\ldots,\varepsilon_{nj},\ldots,\varepsilon_{n,j+nM})^T, \text{ and}$$

$$F = \begin{pmatrix} \left.\dfrac{\partial\hat{\lambda}_1}{\partial\bar{S}_{11}}\right|_{\bar{S}_{**}=\Psi_{**}} & \cdots & \left.\dfrac{\partial\hat{\lambda}_1}{\partial\bar{S}_{n,n+1}}\right|_{\bar{S}_{**}=\Psi_{**}} \\ \vdots & & \vdots \\ \left.\dfrac{\partial\hat{\lambda}_n}{\partial\bar{S}_{11}}\right|_{\bar{S}_{**}=\Psi_{**}} & \cdots & \left.\dfrac{\partial\hat{\lambda}_n}{\partial\bar{S}_{n,n+1}}\right|_{\bar{S}_{**}=\Psi_{**}} \end{pmatrix}. \qquad (5.29)$$

Proof: By a consideration of the vectors ε_{*j} defined above for $j = 0,1,\ldots,M-1$ and from the distributional assumptions that we have made about the random variables ε_{ij} for $i = 1,2,\ldots,n$ and $j = 0,1,\ldots,(n+1)M-1$, we note that the vectors ε_{*j} are independent and identically distributed with zero mean vector and covariance matrix $E(\varepsilon_{*j}\varepsilon_{*j'}^T) = \Sigma \otimes I$ for all $j,\ j' = 0,1,\ldots,M-1$

where $\Sigma = \{\sigma_{ij};\ i = 1,2,\ldots,n,\ j = 1,2,\ldots,n\}$, I is an $(n+1)\times(n+1)$ identity matrix, and $\otimes$ represents the Kronecker or direct product of two square matrices. Let $\bar{\varepsilon}_{*\cdot} = \frac{1}{M}\sum_{j=o}^{M-1} \varepsilon_{*j}$. Then by an application of a form of the multivariate central limit theorem (see Anderson [1958], page 74), we conclude that $\sqrt{M}\,\bar{\varepsilon}_{*\cdot}$ has a limiting multivariate normal distribution with zero mean vector and $\Sigma \otimes I$ as covariance matrix.

Next let us define the vectors

$\bar{S}_{**} = (\bar{S}_{11},\ldots,\bar{S}_{1,n+1},\bar{S}_{21},\ldots,\bar{S}_{2,n+1},\ldots,\bar{S}_{n1},\ldots,\bar{S}_{n,n+1})^T$;

$\zeta_{**} = E(\bar{S}_{**})$; and

$\Psi_{**} = (\Psi_{11},\ldots,\Psi_{1,n+1},\Psi_{21},\ldots,\Psi_{2,n+1},\ldots,\Psi_{n1},\ldots,\Psi_{n,n+1})^T$. Now $\bar{S}_{**} \xrightarrow{p} \Psi_{**}$ as $M\to\infty$, as we demonstrated in the proof of Theorem 5.1. From the definition of our regression model and the above definitions we have:

$$\sqrt{M}(\bar{S}_{**}-\Psi_{**}) = \sqrt{M}(\zeta_{**}-\Psi_{**}) + \sqrt{M}\,\bar{\varepsilon}_{*\cdot} \quad .$$

We want to show that $\lim_{M\to\infty} \sqrt{M}(\zeta_{**}-\Psi_{**}) = 0$, which would imply from the limiting distribution theorem given in Cramér ([1946], page 254) that the limiting distribution of $\sqrt{M}(\bar{S}_{**}-\Psi_{**})$ is the same as the limiting distribution of $\sqrt{M}\,\varepsilon_{*\cdot}$, namely, a multivariate normal distribution with mean vector zero and covariance matrix $\Sigma\otimes I$.

From the definitions of ζ_{iq} and Ψ_{iq} given above and in the proof of Theorem 5.1, we have

$$\sqrt{M}(\zeta_{iq}-\Psi_{iq}) = \sqrt{M} \sum_{k=1}^{n} \left[\frac{1}{M}\alpha_{ik} e^{-\lambda_k(q-1)J} \frac{(1-e^{-\lambda_k J})}{(1-e^{-\lambda_k J/M})} \right.$$

$$\left. - \frac{1}{J}\frac{\alpha_{ik}}{\lambda_k} e^{-\lambda_k(q-1)J}(1-e^{-\lambda_k J}) \right]$$

$$= \sqrt{M} \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k(q-1)J}(1-e^{-\lambda_k J}) \left[\frac{J\lambda_k - M(1-e^{-\lambda_k J/M})}{J\lambda_k(1-e^{-\lambda_k J/M})M} \right]$$

$$= \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k(q-1)J}(1-e^{-\lambda_k J}) \left[\frac{J\lambda_k - M(1-e^{-\lambda_k J/M})}{J\lambda_k(1-e^{-\lambda_k J/M})\sqrt{M}} \right] \to 0 \tag{5.30a}$$

as $M\to\infty$, since by several applications of L'Hôpital's rule

$$\lim_{M\to\infty} \frac{J\lambda_k - M(1-e^{-\lambda_k J/M})}{J\lambda_k(1-e^{-\lambda_k J/M})\sqrt{M}} = 0. \tag{5.30b}$$

Hence $\lim_{M\to\infty} \sqrt{M}(\zeta_{**}-\Psi_{**}) = 0$ and $\sqrt{M}(\bar{S}_{**}-\Psi_{**})$ has a limiting multivariate normal distribution with mean vector zero and covariance matrix $\Sigma \otimes I$. Now let $Z_{**} = M(\bar{S}_{**}-\Psi_{**})$ and expand each member of $\hat{\lambda}$ in a Taylor's series about the point $\bar{S}_{**} = \Psi_{**}$, giving us

$$\hat{\lambda} = \lambda_* + M^{-\frac{1}{2}} F Z_{**}$$

$$+ (2M)^{-1} \left(\sum_{i,q,i',q'} Z_{iq} Z_{i'q'} \frac{\partial^2 \hat{\lambda}_1}{\partial \bar{S}_{iq} \partial \bar{S}_{i'q'}} \Bigg|_{\bar{S}_{**}=\Psi_{**}+M^{-\frac{1}{2}} w_1 Z_{**}} , \ldots , \right.$$

$$\left. \sum_{i,q,i',q'} Z_{iq} Z_{i'q'} \frac{\partial^2 \hat{\lambda}_n}{\partial \bar{S}_{iq} \partial \bar{S}_{i'q'}} \Bigg|_{\bar{S}_{**}=\Psi_{**}+M^{-\frac{1}{2}} w_n Z_{**}} \right)^T \qquad (5.31)$$

where $|w_k| \leq 1$ for $k = 1,2,\ldots,n$. Now let E represent the event that $|\bar{S}_{**}-\Psi_{**}| \leq \delta$ is true for arbitrary $\delta > 0$; $\tilde{E}$ the negation of E; and $E_1 \cap E_2$ the intersection of E_1 and E_2. Then

$$P\{\sqrt{M}(\hat{\lambda}-\lambda_*) \leq x\} = P\{(\sqrt{M}(\hat{\lambda}-\lambda_*) \leq x) \cap E\} + P\{(\sqrt{M}(\hat{\lambda}-\lambda_*) \leq x) \cap \tilde{E}\}. \qquad (5.32a)$$

For any event E_1 we note that

$$P(E_1 \cap \tilde{E}) \leq P(\tilde{E}) \leq \sum_{i,q} P(Z_{iq}^2 \geq M\delta^2), \qquad (5.33a)$$

where the last inequality follows from the definition of E. From a form of Tchebycheff's theorem (see Cramér [1946], page 182) we have:

$$\sum_{i,q} P(Z_{iq}^2 \geq M\delta^2) \leq \sum_{i,q} \frac{E(Z_{iq}^2)}{M\delta^2}$$

$$= \sum_{i,q} \left\{ \frac{\mathrm{Var}(\bar{S}_{iq})}{\delta^2} + \frac{(\zeta_{iq}-\Psi_{iq})^2}{\delta^2} \right\}. \qquad (5.33b)$$

Since $\bar{S}_{iq}$ is the mean of M independent random variables each with the same variance σ_{ii}, then $\mathrm{Var}(\bar{S}_{iq}) = \frac{\sigma_{ii}}{M} \to 0$ as $M\to\infty$. Likewise we know from the proof of Theorem 5.1 that $\zeta_{iq} \to \Psi_{iq}$ as $M\to\infty$. Hence

$$\sum_{i,q} \left\{ \frac{\mathrm{Var}(\bar{S}_{iq})}{\delta^2} + \frac{(\zeta_{iq}-\Psi_{iq})^2}{\delta^2} \right\} = o(1) \text{ or } P(E_1 \cap \tilde{E}) = o(1) \to 0 \quad (5.33c)$$

as $M\to\infty$ from the definition of o(1) (see Cramér [1946], page 122). Therefore

$$P\{\sqrt{M}(\hat{\lambda}-\lambda_*) \leq x\} = P\{(\sqrt{M}(\hat{\lambda}-\lambda_*) \leq x) \cap E\} + o(1). \quad (5.32b)$$

From the hypothesis of our theorem, we are assuming that the functions $\hat{\lambda}_k$ possess continuous second order derivatives of every kind in the neighborhood $|\bar{S}_{**}-\Psi_{**}| \leq \delta$. Therefore we may also conclude that these derivatives are bounded, i.e. there exists a constant C such that

$$\left| \frac{\partial^2 \hat{\lambda}_k}{\partial \bar{S}_{iq} \partial \bar{S}_{i'q'}} \Bigg|_{\bar{S}_{**}=\Psi_{**}+w_k M^{-\frac{1}{2}} Z_{**}} \right| < C \quad (5.34)$$

for all i, q, i′, q′, and k. Hence for each k we have

$$\left| \sum_{i,q,i',q'} M^{-\frac{1}{2}} Z_{iq} Z_{i'q'} \frac{\partial^2 \hat{\lambda}_k}{\partial \bar{S}_{iq} \partial \bar{S}_{i'q'}} \right|_{S_{**}=\Psi_{**}+w_k M^{-\frac{1}{2}} Z_{**}} \Bigg|$$

$$\leq M^{-\frac{1}{2}} C \left(\sum_{i,q} |Z_{iq}| \right)^2 . \tag{5.35}$$

From equation (5.35) we may now write:

$$P\left\{\left[\left(FZ_{**}+ \frac{C}{2\sqrt{M}} \left(\sum_{i,q} |Z_{iq}|\right)^2 \mathit{1}\right) \leq x\right] \cap E\right\} \leq P\left\{ (\sqrt{M}(\hat{\lambda}-\lambda_*) \leq x) \cap E\right\}$$

$$\leq P\left\{\left[\left(FZ_{**}- \frac{C}{2\sqrt{M}} \left(\sum_{i,q} |Z_{iq}|\right)^2 \mathit{1}\right) \leq x\right] \cap E\right\} \tag{5.36}$$

where $\mathit{1}$ is an $n \times 1$ vector with each element equal to one. From equation (5.32b) we may write

$$P\left\{\left[\left(FZ_{**}+ \frac{C}{2\sqrt{M}} \left(\sum_{i,q} |Z_{iq}|\right)^2 \mathit{1}\right) \leq x\right] \cap E\right\} + o(1)$$

$$\leq P\{\sqrt{M}(\hat{\lambda}-\lambda_*) \leq x\} \leq$$

$$P\left\{\left[\left(FZ_{**}- \frac{C}{2\sqrt{M}} \left(\sum_{i,q} |Z_{iq}|\right)^2 \mathit{1}\right) \leq x\right] \cap E\right\} + o(1) . \tag{5.37}$$

We have already found the limiting distribution of Z_{**}, and we want to show that $\frac{C}{2\sqrt{M}} \left(\sum_{i,q} |Z_{iq}|\right)^2 \mathit{1} \xrightarrow{p} 0$ as $M \to \infty$. Consider

$$P\left\{\frac{(\sum_{i,q}(|Z_{iq}|)^2)}{\sqrt{M}} \geq \varepsilon\right\} = P\left\{(\sum_{i,q}(M^{\frac{1}{2}}|\bar{S}_{iq}-\Psi_{iq}|)^2) \geq \varepsilon\sqrt{M}\right\}$$

$$= P\left\{(\sum_{i,q}(|\bar{S}_{iq}-\Psi_{iq}|)^2) \geq \frac{\varepsilon}{\sqrt{M}}\right\} \leq \frac{E\{(\sum_{i,q}(|\bar{S}_{iq}-\Psi_{iq}|)^2)\}\sqrt{M}}{\varepsilon}, \quad (5.38)$$

where the last inequality holds from Tchebycheff's theorem. Next consider

$$\sqrt{M}E(\bar{S}_{iq}-\Psi_{iq})^2 = \sqrt{M}\text{ Var }\bar{S}_{iq} + \sqrt{M}(\zeta_{iq}-\Psi_{iq})^2$$

$$= \sigma_{ii}/\sqrt{M} + \sqrt{M}(\zeta_{iq}-\Psi_{iq})(\zeta_{iq}-\Psi_{iq}) \to 0 \quad (5.39)$$

as $M\to\infty$ by using equation (5.30a). Then consider the term

$$\frac{1}{\sqrt{M}}|Z_{iq}||Z_{i'q'}| = \sqrt{M}|\bar{S}_{iq}-\Psi_{iq}||\bar{S}_{i'q'}-\Psi_{i'q'}|$$

$$= M^{\frac{1}{4}}|\bar{S}_{iq}-\Psi_{iq}|M^{\frac{1}{4}}|\bar{S}_{i'q'}-\Psi_{i'q'}| \quad (5.40)$$

for $i \neq i'$ and $q \neq q'$. By an argument similar to the one used in the proof of Theorem 5.1 to show that $\bar{S}_{iq}-\Psi_{iq} \xrightarrow{p} 0$, we can show that $M^{\frac{1}{4}}(\bar{S}_{iq}-\Psi_{iq}) \xrightarrow{p} 0$. Therefore the expression in equation (5.40) tends in probability to zero as $M\to\infty$. Substituting these results back into equation (5.38) we conclude that $\frac{C}{2\sqrt{M}}(\sum_{i,q}|Z_{iq}|)^2 1$ tends in probability to zero as $M\to\infty$. Therefore the limiting distribution of $\sqrt{M}(\hat{\lambda}-\lambda_*)$ is the same as that of FZ_{**}, i.e. a multivariate normal distribution with mean vector zero and covariance matrix

$F(\Sigma \otimes I)F^T = F\Omega F^T$. This completes the proof of Theorem 5.5.

For the regression model designated as Case II we can use a similar proof in order to prove the following theorem:

Theorem 5.6: Let $\hat{\lambda}'$ represent the $n \times 1$ vector of generalized partial totals estimators for the exponential parameters given in equation (5.10), where each $\hat{\lambda}'_k$, $k = 1,2,\ldots,n$, is a function of the $\bar{S}'_{iq}$, $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$. If each $\hat{\lambda}'_k$ possesses continuous second order derivatives of every kind in the neighborhood $|\bar{S}'_{iq} - \Psi'_{iq}| \leq \delta$ for $\delta > 0$ where $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$, and if the distributional assumptions concerning the random variables ε_{ij} stated at the beginning of Section 5.4 and Theorem 5.1 are satisfied, then the limiting distribution of $\sqrt{M}(\hat{\lambda}' - \lambda'_*)$, where λ'_* represents the vector $\hat{\lambda}'$ with the $\bar{S}'_{iq}$ replaced by the Ψ'_{iq} for all i and q in each of the elements $\hat{\lambda}'_k$, is a multivariate normal distribution with mean vector given by the zero vector and covariance matrix given by $F'\Omega'F'^T$, where $\Omega' = E(\bar{\varepsilon}'_{*}.\bar{\varepsilon}'_{*}.^T)$, $\bar{\varepsilon}'_{*}. = \frac{1}{M}\sum_{j=c}^{M-1} \varepsilon'_{*j}$,

$\varepsilon'_{*j} = (\varepsilon'_{1j}, \varepsilon'_{1,j+M}, \ldots, \varepsilon'_{1,j+nM}, \varepsilon'_{2j}, \ldots, \varepsilon'_{2,j+nM}, \ldots, \varepsilon'_{nj}, \ldots, \varepsilon'_{n,j+nM})^T$,

$\varepsilon'_{ij} = \varepsilon_{ij} - \varepsilon_{i,j+M}$, and

$$F' = \begin{pmatrix} \left.\dfrac{\partial\hat{\lambda}'_1}{\partial\bar{S}'_{11}}\right|_{\bar{S}'_{**}=\Psi'_{**}} & \cdots & \left.\dfrac{\partial\hat{\lambda}'_1}{\partial\bar{S}'_{n,n+1}}\right|_{\bar{S}'_{**}=\Psi'_{**}} \\ \vdots & & \\ \left.\dfrac{\partial\hat{\lambda}'_n}{\partial\bar{S}'_{11}}\right|_{\bar{S}'_{**}=\Psi'_{**}} & \cdots & \left.\dfrac{\partial\hat{\lambda}'_n}{\partial\bar{S}'_{n,n+1}}\right|_{\bar{S}'_{**}=\Psi'_{**}} \end{pmatrix}. \tag{5.41}$$

We now propose to obtain an expression for the asymptotic efficiency of the generalized partial totals estimators of the exponential parameters in our regression models for various distributions of the random variables ε_{ij}. Since the procedure for both Cases I and II are similar, we will present the details only for Case I and quote the results for Case II. Using some of the ideas presented by Kendall and Stuart ([1961], Vol. II, pages 55ff) concerning generalized variances, we will take as our measure of the asymptotic efficiency of our estimators the following ratio:

$$\nu = \lim_{M\to\infty} \left\{ \left| E\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T \right| |\Omega| \right\}^{-1} \tag{5.42}$$

where L represents the likelihood function whose form will be specified and Ω represents the asymptotic covariance matrix of our estimators. Therefore the matrix $E\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T$ will be an $n \times n$ matrix with the $(k,k')^{th}$ element given by $E\left(\frac{\partial \ln L}{\partial \lambda_k}\right)\left(\frac{\partial \ln L}{\partial \lambda_{k'}}\right)$ for $k, k' = 1,2,\ldots,n$. Kendall and Stuart demonstrate that ν is always less than or equal to one when we are considering consistent estimators of the elements of the vector λ, and that $\nu = 1$ for maximum likelihood estimators.

First let us consider for Case I the situation when the joint distribution of the vectors $Y_{*j} = (Y_{1j}, Y_{2j}, \ldots, Y_{nj})^T$, $j = 0,1,\ldots,(n+1)M-1$, is given by:

$$L = \prod_{j=o}^{(n+1)M-1} \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(Y_{*j}-E(Y_{*j}))^T \ \Sigma^{-1}(Y_{*j}-E(Y_{*j}))\right) \tag{5.43}$$

where $\Sigma = E(Y_{*j}-E(Y_{*j}))(Y_{*j}-E(Y_{*j}))^T$ and

$$E(Y_{*j}) = \begin{pmatrix} \sum_{k=1}^{n} \alpha_{1k} e^{-\lambda_k x_j} \\ \sum_{k=1}^{n} \alpha_{2k} e^{-\lambda_k x_j} \\ \vdots \\ \sum_{k=1}^{n} \alpha_{nk} e^{-\lambda_k x_j} \end{pmatrix} . \tag{5.44}$$

Before evaluating the expression $E\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T$ we observe that

$$\ln L = -{}^{n}/_{2} \ln 2\pi - \frac{1}{2}\ln|\Sigma| - \frac{1}{2} \sum_{j=o}^{(n+1)M-1} (Y_{*j}-E(Y_{*j}))^T \Sigma^{-1}(Y_{*j}-E(Y_{*j}))$$

and we need

$$\frac{\partial \ln L}{\partial \lambda} = -\frac{1}{2} \sum_{j=o}^{(n+1)M-1} \frac{\partial}{\partial \lambda} \{(Y_{*j}-E(Y_{*j})^T \Sigma^{-1}(Y_{*j}-E(Y_{*j}))\} .$$

It can be shown from basic matrix theory that since Σ^{-1} is symmetric that

$$\frac{\partial \ln L}{\partial \lambda} = \sum_{j=o}^{(n+1)M-1} D_j \Sigma^{-1}(Y_{*j}-E(Y_{*j}))$$

where the typical $(k,i)^{th}$ element of D_j is given by $\frac{\partial E(Y_{ij})}{\partial \lambda_k}$

From the above results we may now write the following:

$$\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T = \sum_{j=o}^{(n+1)M-1} D_j \Sigma^{-1}(Y_{*j}-E(Y_{*j}))$$

$$\left(\sum_{j=o}^{(n+1)M-1} D_j \Sigma^{-1}(Y_{*j}-E(Y_{*j}))\right)^T$$

$$= \sum_{j=o}^{(n+1)M-1} D_j \Sigma^{-1}(Y_{*j}-E(Y_{*j}))(Y_{*j}-E(Y_{*j}))^T \Sigma^{-1} D_j^T$$

$$+ \sum_{\substack{j,j'=o \\ j\neq j'}}^{(n+1)M-1} D_j \Sigma^{-1}(Y_{*j}-E(Y_{*j}))(Y_{*j'}-E(Y_{*j'}))^T \Sigma^{-1} D_{j'}^T .$$

From the previous assumptions given at the beginning of Section 5.4 concerning the random variables ε_{ij}, we know that

$$\begin{aligned} E(Y_{*j}-E(Y_{*j}))(Y_{*j'}-E(Y_{*j'}))^T &= \Sigma \quad \text{if } j = j' \\ &= 0 \quad \text{if } j \neq j'. \end{aligned}$$

Therefore

$$E\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T = \sum_{j=o}^{(n+1)M-1} D_j \Sigma^{-1} D_j^T \tag{5.45}$$

and from Theorem 5.5 we may now write

$$\nu = \lim_{M\to\infty} \left\{ \left| \sum_{j=o}^{(n+1)M-1} D_j \Sigma^{-1} D_j^T \right| \left| \frac{1}{M} F(\Sigma \otimes I) F^T \right| \right\}^{-1} . \tag{5.46a}$$

Going through similar steps we can determine the following expression for the asymptotic efficiency of the estimators of the exponential parameters for Case II:

$$\nu' = \lim_{M\to\infty} \left\{ \left| \sum_{j=o}^{(n+2)M-1} D_j' \Sigma^{-1} D_j'^{T} \right| \left| \frac{1}{M} F'\Omega' F'^{T} \right| \right\}^{-1} \tag{5.46b}$$

where D_j' is the same as D_j when we use the regression model given by equation (5.10) instead of (5.9) and $F'\Omega' F'^{T}$ is given in Theorem 5.6.

In Chapter 7 where a comparison of the various estimation techniques discussed in this research will be made, the expression for ν will be evaluated for various values of n and various values of the parameters in our regression model.

5.5 Extensions and modifications to the generalized partial totals estimation procedure

5.5.1 Multiple observations at each value of the independent variable

During the discussion in the earlier sections of this chapter concerning the generalized partial totals estimation procedure we have assumed that only one observation was taken for each value of the independent variable. In this section we want to allow for multiple observations to be taken at each value of the independent variable on each regression equation, and we will determine the changes that arise in our generalized estimation procedure. Corresponding to Cases I and II given earlier we have the following new cases:

Case I*: $$Y_{ij\underline{j}} = \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k x_j} + \varepsilon_{ij\underline{j}}$$

for $i = 1,2,\ldots,n$; $j = 0,1,\ldots,(n+1)M-1$; and $\underline{j} = 1,2,\ldots,M_{\underline{i}}$, where the observable random variable $Y_{ij\underline{j}}$ takes on the values denoted by $y_{ij\underline{j}}$.

Case II*: $$Y_{ij\underline{j}} = \alpha_{io} + \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k x_j} + \varepsilon_{ij\underline{j}}$$

for $i = 1,2,\ldots,n$; $j = 0,1,2,\ldots,(n+2)M-1$; and $\underline{j} = 1,2,\ldots,M_{\underline{i}}$, where the observable random variable $Y_{ij\underline{j}}$ takes on the values denoted by $y_{ij\underline{j}}$.

For Case I* we form the new partial totals $\overline{S}^*_{iq}$ defined by

$$\overline{S}^*_{iq} = \frac{1}{MM_{\underline{i}}} \sum_{j=(q-1)M}^{qM-1} \sum_{\underline{j}=1}^{M_{\underline{i}}} y_{ij\underline{j}}$$

where $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$. For the estimation of the exponential parameters of our regression model, we will use exactly the same method as we presented in Section 5.3 with the $\overline{S}^*_{iq}$ substituted in the place of the $\overline{S}_{iq}$. For Case II* we form the new partial totals $\overline{S}^{*\prime}_{iq}$ defined by

$$\overline{S}^{*\prime}_{iq} = \overline{S}^*_{iq} - \overline{S}^*_{i,q+1} \ ,$$

where $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$. For the estimation of the exponential parameters in this case, we will use exactly the same method as we presented for Case II in Section 5.3 with the $\bar{S}^{*\prime}_{iq}$ substituted in the place of the $\bar{S}'_{iq}$.

In order to estimate the linear parameters for our new cases we will proceed just as we did in Section 5.3 for the estimation of the linear parameters in Cases I and II, with

$$\bar{y}_{ij\cdot} = \frac{1}{\underline{M_i}} \sum_{\underline{j}=1}^{\underline{M_i}} y_{ij\underline{j}}$$ replacing the y_{ij} of this earlier section.

In order to conclude anything about the asymptotic properties of these generalized partial totals estimators for these new cases, we need to rephrase our assumptions concerning the random variables $\varepsilon_{ij\underline{j}}$ as follows:

1) For each value of i, j, and $\underline{j}$ we have $E(\varepsilon_{ij\underline{j}}) = 0$ and $E(\varepsilon^2_{ij\underline{j}}) = \sigma_{ii}$ where $0 < \sigma_{ii} < \infty$.

2) For each value of i, i', j, j', $\underline{j}$, and $\underline{j}'$ with $i \neq i'$ and $j \neq j'$, we have $E(\varepsilon_{ij\underline{j}}\varepsilon_{ij'\underline{j}'}) = 0$ and $E(\varepsilon_{ij\underline{j}}\varepsilon_{i'j\underline{j}'}) = \sigma_{ii'}$ where $-\infty < \sigma_{ii'} < \infty$.

3) For each value of $i = 1,2,\ldots,n$ we assume that $\underline{M_i} = \underline{m_i M}$ where $0 < \underline{m_i} < 1$ and $\sum_{i=1}^{n} \underline{m_i} = 1$.

4) The second derivatives of every kind of the estimators of the exponential parameters with respect to the new partial sums $\bar{S}^{*}_{iq}$ or $\bar{S}^{*\prime}_{iq}$ are continuous.

With the above assumptions, we may use the same method of proof as we used in Theorems 5.1 and 5.2, with the $\overline{S}^*_{iq}$'s and $\overline{S}^{*\prime}_{iq}$'s replacing the $\overline{S}_{iq}$'s and $\overline{S}'_{iq}$'s, respectively, to conclude that the estimators of the exponential parameters are consistent as $M\to\infty$ or $\underline{M}\to\infty$. In addition, we may follow a similar procedure to that used by Cornell [1956] and that used in the proof of Theorem 5.5 (or Theorem 5.6) with the $\overline{S}^*_{iq}$'s (or $\overline{S}^{*\prime}_{iq}$'s) replacing the $\overline{S}_{iq}$'s (or $\overline{S}'_{iq}$'s) to demonstrate that $\sqrt{M\underline{M}}(\hat{\lambda}^*-\lambda^*_*)$ $\left(\text{or } \sqrt{M\underline{M}}(\hat{\lambda}^{*\prime}-\lambda^{*\prime}_*)\right)$ has a limiting multivariate normal distribution as $\underline{M}\to\infty$ or $M\to\infty$, where $\hat{\lambda}^*$ and λ^*_* (or $\hat{\lambda}^{*\prime}$ and $\lambda^{*\prime}_*$) have similar definitions to $\hat{\lambda}$ and λ_*(or $\hat{\lambda}'$and λ'_*) given in Theorem 5.5 (or Theorem 5.6).

5.5.2 Some justification for the grouping of our partial totals and some modifications for unequally spaced values of x_j

During the development and evaluation of our generalized partial totals estimation procedure, no reason was given for the particular grouping of our observations in order to find the partial totals of S_{iq}, which, in turn, would give us our estimators of the exponential parameters. In this discussion we will demonstrate that some other forms of grouping our observations do not lead to estimators with the desirable properties that our estimators have. Most of the results presented here follow directly from the work by Cornell [1956], but will be presented here for completeness. Also, although the following results may be obviously extended to Cases II, I*, and II*, we will give the results only for Case I.

One possible alternative to the grouping of our observations would be to divide the range of our independent variable into M equal portions, taking the first observation of each segment to go into the first partial total, taking the second observation of each segment to go into the second partial total, etc. Our new partial totals would then be written as:

$$S^A_{iq} = \sum_{j=o}^{M-1} y_{i,j(n+1)+q-1}$$

for $i = 1,2,\ldots,n$ and $q = 1,2,\ldots,n+1$. Corresponding to equation (5.11) we have

$$\Sigma^A_{iq} = \sum_{j=o}^{M-1} \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k h(j(n+1)+q-1)}$$

$$= \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k h(q-1)} \frac{(1-e^{-\lambda_k h(n+1)M})}{(1-e^{-\lambda_k h(n+1)})} .$$

Going through the same basic steps that we used in Sections 5.2 and 5.3, we arrive at the following equation corresponding to equation (5.12):

$$\Lambda^A_n \Sigma^A_{i1} - \Lambda^A_{n-1} \Sigma^A_{i2} + \Lambda^A_{n-2} \Sigma^A_{i3} - \ldots + (-1)^{n+1} \Lambda^A_o \Sigma^A_{i,n+1} = 0$$

for $i = 1,2,\ldots,n$, where Λ^A_r, $r = 1,2,\ldots,n$, are the elementary symmetric functions of $e^{-\lambda_k h}$. In addition, we define $\Lambda_o^A = 1$. By substituting the S_{iq}^A's in the place of the Σ_{iq}^A's, it follows that the solutions for the exponential parameters may be easily obtained in a similar manner to that used in Section 5.3.

In order to investigate some of the properties of these estimators we will make the same assumptions concerning the random variables ε_{ij} as stated at the beginning of Section 5.4. Next let us consider:

$$\bar{S}_{iq}^A = \frac{1}{M} S_{iq}^A = \frac{1}{M} \sum_{j=o}^{M-1} \sum_{k=1}^{n} \alpha_{ik} e^{-\lambda_k h(j(n+1)+q-1)} + \frac{1}{M} \sum_{j=o}^{M-1} \varepsilon_{i,j(n+1)+q-1}$$

$$= \frac{1}{M} \Sigma_{iq}^A + \frac{1}{M} \sum_{j=o}^{M-1} \varepsilon_{i,j(n+1)+q-1} \ .$$

By the same reasoning as we used in the proof of Theorem 5.1, we can show that $\frac{1}{M} \sum_{j=o}^{M-1} \varepsilon_{i,j(n+1)+q-1} \xrightarrow{p} 0$ as $M \to \infty$. By using the same assumptions as given in Theorem 5.1 we can show that

$$\lim_{M\to\infty} \frac{1}{J} \sum_{k=1}^{n} \sum_{j=o}^{M-1} \alpha_{ik} e^{-\lambda_k J(n+1)j/M} e^{-\lambda_k J(q-1)/M} \left(J/_M\right)$$

$$= \frac{1}{J} \sum_{k=1}^{n} \alpha_{ik} \int_o^J e^{-\lambda_k (n+1)t} dt = \Psi_{iq}^A,$$

which is independent of q. Hence $\bar{S}_{iq}^A \xrightarrow{p} \Psi_{iq}^A$, and by using the same line of reasoning as presented by Cornell [1956] we conclude that the matrix R^A evaluated at the point $\bar{S}_{iq}^A = \Psi_{iq}^A$ will be singular, and therefore our estimators of the exponential parameters will $\vec{p}$ ∞ as $M \to \infty$. That is, these estimators will not be consistent.

Another alternative method of grouping could be applied when M is a multiple of (n+1), in which case we could divide our domain into $(n+1)^2$ equal groups each containing $\frac{M}{n+1}$ observations. Then in the place of Σ_{iq} we will use

$$\Sigma_{iq}^A = \sum_{j=(q-1)\frac{M}{n+1}}^{q\frac{M}{n+1}-1} E(Y_{ij}) + \sum_{j=(q+n)\frac{M}{n+1}}^{(q+n+1)\frac{M}{n+1}-1} E(Y_{ij}) + \ldots + \sum_{j=(q+n(n+1))\frac{M}{n+1}}^{(q+n(n+1)+1)\frac{M}{n+1}-1} E(Y_{ij})$$

where $i = 1,2,\ldots,n$; $q = 1,2,\ldots,n+1$; and the superscript A indicates an alternative grouping. With this alternative grouping we can show that the new partial totals satisfy the system of equations given by (5.12) with the Σ_{iq}^A substituted for the Σ_{iq}. We may now use this new system of equations to solve for the elementary symmetric functions of $e^{-\lambda_1 d\frac{M}{n+1}}, \ldots, e^{-\lambda_n d\frac{M}{n+1}}$. In order to find our estimators of these elementary symmetric functions we will use the partial totals $\bar{S}_{iq}^A$ in place of Σ_{iq}^A where

$$\bar{S}^{A}_{iq} = \frac{n+1}{M}\left\{ \sum_{j=(q-1)\frac{M}{n+1}}^{q\frac{M}{n+1}-1} y_{ij} + \sum_{j=(q+n)\frac{M}{n+1}}^{(q+n+1)\frac{M}{n+1}-1} y_{ij} + \ldots + \sum_{j=(q+n(n+1))\frac{M}{n+1}}^{(q+n(n+1)+1)\frac{M}{n+1}-1} y_{ij} \right\}.$$

From the results given in Section 5.4 and the results derived by Cornell [1956] we can show for this alternative grouping that the determinant of the matrix R^A, which is the same as the matrix R in Section 5.4 with $\bar{S}^{A}_{iq}$ used in place of $\bar{S}_{iq}$, evaluated at $\bar{S}^{A}_{iq} = \phi^{A}_{iq}$ is equal to zero, where $\bar{S}^{A}_{iq} \xrightarrow{p} \phi^{A}_{iq}$. Hence we see that there are alternative groupings for the partial totals which lead to an estimation procedure similar to that developed in Section 5.3, but these alternative groupings do not have some of the desirable properties of the generalized procedure presented in Section 5.3.

We now want to propose some modifications that will be concerned with the assumption about the values of x_j in equations (5.9) and (5.10) being equally spaced, and with the assumption about an equal number of values of x_j being taken for each partial total S_{iq}. For the particular situations arising when these assumptions are not satisfied, we may think of the Σ_{iq}'s as approximations to areas under the curve found by plotting $E(Y_i(x))$ against the independent variable x. With this interpretation in mind, we will suggest the following modifications to the generalized partial totals estimation procedure when some of our observations do not satisfy all of the basic assumptions:

1) Divide the domain of the independent variable into the desired number of intervals, each interval being of the same length.

2) Add the observations in each of these new intervals together, weighting each observation y_{ij} by the following:

$$\frac{x_{j+1}-x_j}{2} + \frac{x_j-x_{j-1}}{2} = \frac{x_{j+1}-x_{j-1}}{2} .$$

3) Divide these new partial totals by the sum of the weights, and substitute these weighted averages in the place of the $\overline{S}_{iq}$'s which appeared in Section 5.3.

In order for the limiting properties of the estimators to still hold, we must continue to assume that the domain of the independent variable for each partial total is constant, and the number of observations for each partial total becomes large. The above modifications are useful for the estimation of the exponential parameters. The estimation of the linear parameters, given the exponential estimates, would remain unchanged.

VI. A GENERALIZATION OF SPEARMAN ESTIMATION

6.1 Introduction

In this chapter we will develop and discuss some of the properties of an estimation procedure that may be applied to the estimation of the exponential parameters of a member of the class of regression models given by equation (1.1). This estimation procedure will be based upon a generalization of the Spearman estimation technique as presented by Johnson and Brcwn [1961]. Therefore in Section 6.2 we will present the estimation technique for the case when $n = m = 1$ in equation (1.1), which is the case considered by Johnson and Brown. It is given here for completeness. In Section 6.3 the generalization to this estimation procedure will be presented, and finally in Section 6.4 certain properties of these estimators will be presented. Since the particular regression models that we will be considering are motivated by the equations that arise in describing tracer experiments, there will be certain restrictions placed upon the linear coefficients of our model, e.g. the sum of these coefficients must be equal to a constant. As the various steps of the estimation procedure are presented we will specify these restrictions.

6.2 Single equation Spearman estimation

The results in this section are contained in the work by Johnson and Brown [1961] and Cornell [1965]. However, we will repeat some of the results here since some of them will be needed for the generalized estimation procedure to be presented in the next section. The particular regression model that we are considering now is given by:

$$Y_j = 1 - e^{-\lambda x_j} + \varepsilon_j \ . \tag{6.1a}$$

It is possible to think of this as one of the regression equations arising from a two-compartment mammillary or catenary system where the observations represent the proportion of radioactive tracer present in a compartment at a particular time, as Section 3.1 demonstrates. Since the sum of the expected values of the observed random variables will be fixed for this case, there is only one independent regression equation which is given by (6.1a). We will drop the subscript i as we did in some of the earlier chapters, since for this discussion $i = 1$ only. In the last chapter we assumed that the values of the independent variable were taken such that $x_{j+1} - x_j = h$ for all j where $h > 0$ and h is independent of j. Since we are assuming that the expected values of the observed random variables are following an exponential function, it would appear more reasonable in fitting our observations to a particular function to take most of our observations in that region of the independent variable where our function is changing the most.

Hence in this chapter we will take our independent variable to be of the form $x_j = e^{z_j}$ where $z_j = z_o + jd$ for $j = 0, \pm 1, \pm 2, \ldots, \pm M'$ when the number of observations is odd, and $z_j = z_o + d(j + \frac{1}{2})$ for $j = 0, \pm 1, \pm 2, \ldots, \pm (M' - 1), - M'$ when the number of observations is even, i.e. the values of z_j are taken to be equally spaced. Hence our regression model becomes

$$Y_j = 1 - \exp(-\lambda e^{z_j}) + \varepsilon_j, \tag{6.1b}$$

for $j = 0, \pm 1, \pm 2, \ldots, \pm (M'-1)$, and $\pm M'$ or $- M'$ depending upon whether an odd or even number of observations has been taken. When we want to consider the expected value of the observed random variable as a continuous function of the independent variable, we will denote this by $E(Y(z)) = 1 - \exp(-\lambda e^{z})$ where $-\infty < z < \infty$. This last statement implies that we are assuming that $E(\varepsilon_j) = 0$ for all j, where ε_j appears in equation (6.1b).

The first step in this estimation procedure will be to evaluate the following integral:

$$\mu^{(1)} = \int_{-\infty}^{\infty} z \, dE(Y(z)) = \int_{-\infty}^{\infty} z \lambda e^{z} \exp(-\lambda e^{z}) \, dz = -\gamma - \ln\lambda \tag{6.2}$$

where γ is Euler's constant. Solving the above equation for λ we find

$$\lambda = e^{-\gamma} e^{-\mu^{(1)}}. \tag{6.3a}$$

The next step in this estimation scheme will be to propose an estimator of $\mu^{(1)}$, denoted by $\hat{\mu}^{(1)}$, which will be substituted into (6.3a) to give us the following estimator of λ:

$$\hat{\lambda} = e^{-\gamma}\, e^{-\hat{\mu}^{(1)}} . \tag{6.3b}$$

The estimator for $\mu^{(1)}$ is given by the following expression:

$$\hat{\mu}^{(1)} = \sum_{j=-M'}^{M''-1} \left(\frac{z_j + z_{j+1}}{2}\right) \Delta y_j \ , \tag{6.4}$$

where $\Delta y_j = y_{j+1} - y_j$, y_j represents the observation for the value of the independent variable equal to z_j, and where $M'' = M'$ for an odd number of observations but $M'' = M' - 1$ for an even number of observations.

After proposing the above estimator for $\mu^{(1)}$, Johnson and Brown [1961] investigate some of the properties of $\hat{\mu}^{(1)}$ under the following assumptions:

1) The observed random variables are independent binomially distributed random variables.

2) The value of M' is assumed to be large enough so that we can take $y_{-M'} = E(Y(-\infty)) = 0$ and $y_{M''} = E(Y(+\infty)) = 1$, where M'' has been defined above. Under the above assumptions Johnson and Brown demonstrate that $\hat{\mu}^{(1)}$ is approximately unbiased and that the variance of $\hat{\mu}^{(1)}$ is approximately equal to $\frac{d \ln 2}{n^*}$ where n^* is the number of observations taken at each value of the independent variable. These authors also demonstrate that the asymptotic

efficiency of $\hat{\lambda}$ as M' becomes large is 88 per cent.

Cornell [1965] demonstrates the similarities between the Spearman estimation procedure proposed by Johnson and Brown and the estimation procedure proposed by Fisher [1921] for the model given by the single exponential equation.

6.3 Generalization of Spearman estimation

In this section we will show how the method of Spearman estimation discussed in Section 6.2 may be generalized to estimate the exponential parameters in another regression model that is a particular member of the class of regression models given by equation (1.1) for $n > 1$. The particular regression model that we are interested in may be specified by:

$$Y_i(z) = E(Y_i(z)) + \varepsilon_i \tag{6.5a}$$

for $i = 1,2,\ldots,n+1$ and $-\infty < z < \infty$. More specifically, we will take our regression model as:

$$E(Y_i(z)) = \alpha_{i1} \exp(-\lambda_1 e^z) + \alpha_{i2} \exp(-\lambda_2 e^z) + \ldots$$
$$+ \alpha_{in} \exp(-\lambda_n e^z) + \alpha_{i,\, n+1} \tag{6.5b}$$

for $i = 1,2,\ldots,n+1$, where the following conditions are imposed:

1) $E(Y_1(-\infty)) = E(Y_2(+\infty)) = 1.$

2) $E(Y_1(+\infty)) = E(Y_2(-\infty)) = E(Y_3(\pm\infty)) = E(Y_4(\pm\infty)) = \ldots = E(Y_n(\pm\infty))$
$= E(Y_{n+1}(\pm\infty)) = 0.$

3) $\sum_{i=1}^{n+1} E(Y_i(z)) = 1$ for all z and $0 \leq E(Y_i(z)) \leq 1$ for all i and z.

With the above conditions we may find the following relations to be satisfied by the coefficients of (6.5b):

1) $\alpha_{1,n+1} = 0$; $\alpha_{1n} = 1 - \alpha_{11} - \alpha_{12} - \dots - \alpha_{1,n-1}$,

2) $\alpha_{2,n+1} = 1$; $\alpha_{2n} = -(\alpha_{21}+\alpha_{22}+\dots+\alpha_{2,n-1}+\alpha_{2,n+1})$

$$= -(1+\alpha_{21}+\dots+\alpha_{2,n-1});$$

3) $\alpha_{i,n+1} = 0$ for $i = 3,4,\dots,n+1$,

4) $\alpha_{in} = -(\alpha_{i1}+\alpha_{i2}+\dots+\alpha_{i,n-1})$ for $i = 3,4,\dots,n$,

5) $\alpha_{n+1,k} = -\sum_{i=1}^{n} \alpha_{ik}$ for $k = 1,2,\dots,n$.

It is easily seen by reference to the theorems of Section 3.1 that the regression model given by (6.5b) with the above conditions may be used to describe an (n+1)-compartment catenary or mammillary model where a fixed amount of tracer material is injected into the first compartment of the system and is allowed to accumulate in the second compartment of the system. The observations would represent the proportion of tracer material present in a particular compartment for a particular value of the independent variable. By a comparison of the results from the theorems in Section 3.1 with equation (6.5b), it is noted that we have taken $x = e^z$. The reasons for doing this are the same as those given in Section 6.2 for the simple exponential model. With the above conditions on the coefficients of our exponential terms in the regression model, we may now write our expressions for the expected values as:

$$E(Y_1(z)) = \alpha_{11} \exp(-\lambda_1 e^z) + \alpha_{12} \exp(-\lambda_2 e^z) + \dots$$

$$+ \alpha_{1,n-1} \exp(-\lambda_{n-1} e^z) + (1-\alpha_{11}-\alpha_{12}- \dots -\alpha_{1,n-1}) \exp(-\lambda_n e^z);$$

$$E(Y_2(z)) = 1 + \alpha_{21} \exp(-\lambda_1 e^z) + \alpha_{22} \exp(-\lambda_2 e^z) + \dots$$

$$+ \alpha_{2,n-1} \exp(-\lambda_{n-1} e^z) - (1+\alpha_{21}+ \dots +\alpha_{2,n-1}) \exp(-\lambda_n e^z);$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$E(Y_n(z)) = \alpha_{n1} \exp(-\lambda_1 e^z) + \alpha_{n2} \exp(-\lambda_2 e^z) + \dots$$

$$+ \alpha_{n,n-1} \exp(-\lambda_{n-1} e^z) - (\alpha_{n1}+\alpha_{n2}+ \dots + \alpha_{n,n-1}) \exp(-\lambda_n e^z);$$

$$E(Y_{n+1}(z)) = - (\alpha_{11}+\alpha_{21}+ \dots +\alpha_{n1}) \exp(-\lambda_1 e^z)$$

$$- (\alpha_{12}+\alpha_{22}+ \dots +\alpha_{n2}) \exp(-\lambda_2 e^z) - \dots$$

$$- (\alpha_{1,n-1}+\alpha_{2,n-1}+ \dots +\alpha_{n,n-1}) \exp(-\lambda_{n-1} e^z)$$

$$+ (\alpha_{11}+ \dots +\alpha_{n,n-1}) \exp(-\lambda_n e^z). \tag{6.5c}$$

From the assumption that $\sum_{i=1}^{n+1} E(Y_i(z)) = 1$ for all z, we note that only n of the equations given by (6.5b) or (6.5c) are independent. Therefore, without loss of generality, we will work with the first n equations of this set. The basic steps of the generalized estimation procedure may be outlined as follows:

Step 1: A linear combination of the n elementary symmetric functions of $\ln\lambda_1$, $\ln\lambda_2,\dots,\ln\lambda_n$ is derived for each of the n independent regression equations.

Step 2: The above linear system of equations is solved for the elementary symmetric functions.

Step 3: Expressions for $\ln\lambda_1$, $\ln\lambda_2,\ldots,\ln\lambda_n$ are obtained by finding the roots of an n^{th} order polynomial equation.

Step 4: We use the relations $\lambda_1 = e^{\ln\lambda_1},\ldots,\lambda_n = e^{\ln\lambda_n}$ to obtain expressions for $\lambda_1,\ldots,\lambda_n$, which are functions of $E(Y_1(z)),\ldots,$ $E(Y_n(z))$.

Step 5: We obtain our estimators of $\lambda_1,\lambda_2,\ldots,\lambda_n$ by approximating the functions of $E(Y_1(z))$, $E(Y_2(z)),\ldots,E(Y_n(z))$ by functions of our observations y_{ij}, where y_{ij} is an observation on the i^{th} equation for the value of the independent variable equal to z_j.

Before we can do the first step of our estimation procedure, we must consider two different cases which cover each of the equations in (6.5c) and which are given by:

Case 1: The regression equation may be reduced to the form:

$$E(Y_i(z)) = \sum_{k=1}^{n-1} \alpha_{ik} \exp(-\lambda_k e^z) + (1-\alpha_{i1}-\ldots-\alpha_{i,n-1}) \exp(-\lambda_n e^z). \tag{6.6a}$$

Case 2: The regression equation may be reduced to the form:

$$E(Y_i(z)) = \alpha_{i1} \exp(-\lambda_1 e^z) +\ldots+ \alpha_{i,n-1} \exp(-\lambda_{n-1} e^z) - (\alpha_{i1}+\ldots+\alpha_{i,n-1}) \exp(-\lambda_n e^z). \tag{6.6b}$$

Since the subscript i will not be needed in the following derivation, we will drop this subscript during the following discussion.

First we will consider Case 1 where the regression equation may be described by:

$$E(Y(z)) = \sum_{k=1}^{n-1} \alpha_k \exp(-\lambda_k e^z)$$

$$+ (1-\alpha_1-\ldots-\alpha_{n-1}) \exp(-\lambda_n e^z). \tag{6.7}$$

Consider the following integral:

$$\mu^{(k')} = \int_{-\infty}^{\infty} z^{k'} dE(Y(z)) = \int_{-\infty}^{\infty} z^{k'} d\left\{\sum_{k=1}^{n-1} \alpha_k \exp(-\lambda_k e^z)\right.$$

$$\left. + (1-\alpha_1-\ldots-\alpha_{n-1}) \exp(-\lambda_n e^z)\right\}$$

$$= -\int_{-\infty}^{\infty} z^{k'} \left\{\sum_{k=1}^{n-1} \alpha_k \lambda_k e^z \exp(-\lambda_k e^z)\right.$$

$$\left. + (1-\alpha_1-\ldots-\alpha_{n-1}) \lambda_n e^z \exp(-\lambda_n e^z)\right\} dz \tag{6.8}$$

for $k' = 1,2,\ldots,n-1$. A typical term in (6.8) would be of the form

$$-\int_{-\infty}^{\infty} \alpha z^{k'} \lambda e^z \exp(-\lambda e^z) dz$$

which reduces to

$$-\alpha \int_{0}^{\infty} (\ln t - \ln\lambda)^{k'} e^{-t} dt = -\alpha \left\{ I_{k'} - \binom{k'}{1} \ln\lambda I_{k'-1}\right.$$

$$\left. + \binom{k'}{2} (\ln\lambda)^2 I_{k'-2} + \ldots + \binom{k'}{k'-1} (-\ln\lambda)^{k'-1} I_1 + (-\ln\lambda)^{k'} I_o \right\}$$

by the substitution $t = \lambda e^z$ where $I_{k''} = \int_0^{\infty} (\ln t)^{k''} e^{-t} dt$ for $k'' = 0,1,\ldots,k'$. By substituting the above expressions for a typical term of (6.8) back into (6.8), we obtain the following system of equations:

$$\sum_{k=1}^{n-1}\left\{\binom{k'}{1}I_{k'-1}(\ln\lambda_k-\ln\lambda_n)-\binom{k'}{2}I_{k'-2}((\ln\lambda_k)^2-(\ln\lambda_n)^2)+\dots\right.$$

$$+(-1)^{k'-2}\binom{k'}{k'-1}I_1((\ln\lambda_k)^{k'-1}-(\ln\lambda_n)^{k'-1})$$

$$\left.+(-1)^{k'-1}((\ln\lambda_k)^{k'}-(\ln\lambda_n)^{k'})\right\}\alpha_k=$$

$$\mu^{(k')}+I_{k'}-\binom{k'}{1}I_{k'-1}\ln\lambda_n+\binom{k'}{2}I_{k'-2}(\ln\lambda_n)^2+\dots$$

$$+\binom{k'}{k'-1}I_1(-\ln\lambda_n)^{k'-1}+(-\ln\lambda_n)^{k'} \tag{6.9a}$$

for $k' = 1,2,\dots,n-1$. In order to simplify the notation in (6.9a), we make the substitutions $l_k = \ln\lambda_k$ and $\alpha_k' = \alpha_k(l_k-l_n)$ for $k = 1,2,\dots,n$. Equation (6.9a) becomes

$$\sum_{k=1}^{n-1}\left\{\binom{k'}{1}I_{k'-1}-\binom{k'}{2}I_{k'-2}\left(\frac{l_k^2-l_n^2}{l_k-l_n}\right)+\binom{k'}{3}I_{k'-3}\left(\frac{l_k^3-l_n^3}{l_k-l_n}\right)-\right.$$

$$\dots+(-1)^{k'-2}\binom{k'}{k'-1}I_1\left(\frac{l_k^{k'-1}-l_n^{k'-1}}{l_k-l_n}\right)$$

$$\left.+(-1)^{k'-1}\left(\frac{l_k^{k'}-l_n^{k'}}{l_k-l_n}\right)\right\}\alpha_k'=$$

$$\mu^{(k')}+I_{k'}+\binom{k'}{1}I_{k'-1}(-l_n)+\binom{k'}{2}I_{k'-2}(-l_n)^2+\dots$$

$$+\binom{k'}{k'-1}I_1(-l_n)^{k'-1}+(-l_n)^{k'} \tag{6.9b}$$

for $k' = 1,2,\dots,n-1$. A small table of values of the function $I_{k'}$ for various values of k' is given as follows:

k'	$I_{k'}$
0	1.00000000
1	-0.57721566
2	1.97811199
3	-0.63664683
4	12.45795881
5	-80.84065721
6	486.79308438

At this point we will use matrix notation so that we may express all of the equations in (6.9b) in a single expression. Let V be an $(n-1) \times 1$ vector with k'^{th} element given by

$$\mu^{(k')} + I_{k'} + \binom{k'}{1} I_{k'-1}(-1_n) + \binom{k'}{2} I_{k'-2}(-1_n)^2 + \ldots$$

$$+ \binom{k'}{k'-1} I_1 (-1_n)^{k'-1} + (-1_n)^{k'}. \qquad (6.10)$$

Let α' be an $(n-1) \times 1$ vector with elements: $\alpha_1', \alpha_2', \ldots, \alpha_{n-1}'$. Let L' be an $(n-1) \times (n-1)$ matrix with $(k',k)^{th}$ element given by

$$\sum_{k''=1}^{k'} \binom{k'}{k''} (-1)^{k''-1} I_{k'-k''} \left(\frac{1_k^{k''} - 1_n^{k''}}{1_k - 1_n} \right). \qquad (6.11)$$

Using the above defined vectors and matrices we may write (6.9b) as

$$L' \alpha' = V. \qquad (6.9c)$$

For the particular case when $n = 4$, the matrices L', α', and V may be written as

$$L' = \begin{pmatrix} 1 & 1 & 1 \\ 2I_1 - (l_1+l_4) & 2I_1 - (l_2+l_4) & 2I_1 - (l_3+l_4) \\ 3I_2 - 3I_1(l_1+l_4) + (l_1^2+l_1l_4+l_4^2) & 3I_2 - 3I_1(l_2+l_4) + (l_2^2+l_2l_4+l_4^2) & 3I_2 - 3I_1(l_3+l_4) + (l_3^2+l_3l_4+l_4^2) \end{pmatrix},$$

$$\alpha' = \begin{pmatrix} \alpha_1' \\ \alpha_2' \\ \alpha_3' \end{pmatrix}, \text{ and } V = \begin{pmatrix} \mu^{(1)} + I_1 - l_4 \\ \mu^{(2)} + I_2 - 2I_1l_4 + l_4^2 \\ \mu^{(3)} + I_3 - 3I_2l_4 + 3I_1l_4^2 - l_4^3 \end{pmatrix}.$$

Next carry out the following set of elementary row operations on the matrices L' and V:

1) Multiply the first row by $- 3I_2$ and add to the third row.
2) Multiply the first row by $- 2I_1$ and add to the second row.
3) Multiply the second row by $- 3I_1$ and add to the third row.
4) Multiply the second row by l_4 and add to the third row.

5) Multiply the first row by l_4 and add to the second row.

After the above operations the matrix L´ reduces to

$$\begin{pmatrix} 1 & 1 & 1 \\ -l_1 & -l_2 & -l_3 \\ l_1^2 & l_2^2 & l_3^2 \end{pmatrix},$$

and the matrix V reduces to

$$\begin{pmatrix} K_1 - l_4 \\ K_2 + K_1 l_4 \\ K_3 + K_2 l_4 \end{pmatrix},$$

where $K_1 = \mu^{(1)} + I_1$; $K_2 = \mu^{(2)} + I_2 - 2I_1K_1$; and $K_3 = \mu^{(3)} + I_3 - 3I_2K_1 - 3I_1K_2$. The performance of the above elementary row operations is equivalent to multiplying both sides of (6.9c) by the triangular matrix

$$T = \begin{pmatrix} 1 & 0 & 0 \\ -2I_1 + l_4 & 1 & 0 \\ 6I_1^2 - 3I_2 - 2I_1l_4 & -3I_1 + l_4 & 1 \end{pmatrix}.$$

By the use of an inductive argument it can be shown in general as well as for n = 4 that there exists an (n-1) × (n-1) triangular matrix T such that

$$TL = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ -l_1 & -l_2 & -l_3 & \cdots & -l_{n-1} \\ l_1^2 & l_2^2 & l_3^2 & \cdots & l_{n-1}^2 \\ -l_1^3 & -l_2^3 & -l_3^3 & \cdots & -l_{n-1}^3 \\ \vdots & \vdots & \vdots & & \vdots \\ (-l_1)^{n-2} & (-l_2)^{n-2} & (-l_3)^{n-2} & \cdots & (-l_{n-1})^{n-2} \end{pmatrix} \tag{6.12}$$

and

$$TV = \begin{pmatrix} K_1 - l_n \\ K_2 + K_1 l_n \\ K_3 + K_2 l_n \\ \vdots \\ K_{n-1} + K_{n-2} l_n \end{pmatrix} \tag{6.13}$$

where

$$K_1 = \mu^{(1)} + I_1$$

$$K_2 = \mu^{(2)} + I_2 - 2I_1K_1$$

$$K_3 = \mu^{(3)} + I_3 - 3I_2K_1 - 3I_1K_2$$

$$\vdots$$

$$K_{n-1} = \mu^{(n-1)} + I_{n-1} - \binom{n-1}{1} I_{n-2}K_1 - \binom{n-1}{2} I_{n-3}K_2 - \cdots$$

$$- \binom{n-1}{n-2} I_1 K_{n-2}. \qquad (6.14)$$

Using the above notation, the solution for α' becomes

$$\alpha' = (TL')^{-1}\, TV. \qquad (6.15)$$

We note that the matrix TL' is a Vandermonde matrix; therefore,

$$|TL'| = \prod_{k < k'} (l_k - l_{k'}) \neq 0 \qquad (6.16)$$

and $(TL')^{-1}$ exists since $l_k \neq l_{k'}$ for $k \neq k'$ because we have assumed that $\lambda_k \neq \lambda_{k'}$ for $k \neq k'$.

In order to simplify some of the above notation let $D = TL'$ and let $D^{(k)}$ be an $(n-2) \times (n-2)$ matrix derived from D by deleting the k^{th} column and $(n-1)^{st}$ (or last) row of D. Let C^* represent the set $\{l_1, l_2, \ldots, l_{n-1}\}$ and let $C^*_{(k)} = C^* - \{l_k\}$. Then from Aitken ([1949], page 118) or Cornell ([1956], page 27) the $(k', k)^{th}$ element of the matrix D^{-1} is given by

$$d^{k'k} = (-1)^{n+k'-1} \frac{|D^{(k')}|}{|D|} \Lambda_{n-k-1}^{(C^*_{(k')})} \tag{6.17}$$

where $\Lambda_{n-k-1}^{(C^*_{(k')})}$ is the $(n-k-1)^{th}$ elementary symmetric function of the elements from the set $C^*_{(k')}$. Using equation (6.15), the k'^{th} element of α' is given as:

$$\alpha'_{k'} = \sum_{k=1}^{n-1} (-1)^{n+k'-1} \frac{|D^{(k')}|}{|D|} \Lambda_{n-k-1}^{(C^*_{(k')})} (K_k + K_{k-1} 1_n) \tag{6.18a}$$

where $K_o = -1$.

If we define $\Lambda_o^{(C^*_{(k')})} = 1$ and use the relation

$$1_n \Lambda_k^{(C^*_{(k')})} = \Lambda_{k+1}^{(C_{(k')})} - \Lambda_{k+1}^{(C^*_{(k')})},$$

where $C_{(k')} = C^*_{(k')} + \{1_n\}$, then

$$\alpha'_{k'} = (-1)^{n+k'-1} \frac{|D^{(k')}|}{|D|} \sum_{k=o}^{n-1} K_{n-k-1} \Lambda_k^{(C_{(k')})}. \tag{6.18b}$$

The next step will be to substitute the vector α' into the following equation:

$$\sum_{k=1}^{n-1} \sum_{k''=1}^{n} (-1)^{k''-1} \binom{n}{k''} I_{n-k''} \left(\frac{1_k^{k''} - 1_n^{k''}}{1_k - 1_n} \right) \alpha'_k$$

$$= \mu^{(n)} + I_n + \sum_{k''=1}^{n} \binom{n}{k''} I_{n-k''} (-1_n)^{k''}. \tag{6.19}$$

With this substitution the following equation is satisfied:

$$\begin{vmatrix} L' & -V \\ 1'^T & -v \end{vmatrix} = 0 , \tag{6.20}$$

where the matrices L′ and V have already been defined earlier, and $1'$ is an $(n-1) \times 1$ vector with k^{th} element

$$\binom{n}{1} I_{n-1} - \binom{n}{2} I_{n-2} \left(\frac{1_k^2 - 1_n^2}{1_k - 1_n}\right) + \binom{n}{3} I_{n-3} \left(\frac{1_k^3 - 1_n^3}{1_k - 1_n}\right) + \ldots$$
$$+ (-1)^{n-2} \binom{n}{n-1} I_1 \left(\frac{1_k^{n-1} - 1_n^{n-1}}{1_k - 1_n}\right) + (-1)^{n-1} \left(\frac{1_k^n - 1_n^n}{1_k - 1_n}\right) \tag{6.21}$$

and

$$v = \mu^{(n)} + I_n + \sum_{k''=1}^{n} \binom{n}{k''} I_{n-k''} (-1_n)^{k''} . \tag{6.22}$$

It should be noted that equation (6.19) is merely equation (6.9a) with $k' = n$.

For the case $n = 4$ the matrix $\begin{pmatrix} L' & -V \\ 1'^T & -v \end{pmatrix}$ is equal to the display given on the following page:

$$\begin{pmatrix} 1 & 1 & 1 & \mu^{(1)} + I_1 - l_4 \\ \sum_{k''=1}^{2}\binom{2}{k''}(-1)^{k''-1}I_{2-k''}\left(\frac{l_1^{k''} - l_4^{k''}}{l_1 - l_4}\right) & \sum_{k''=1}^{2}\binom{2}{k''}(-1)^{k''-1}I_{2-k''}\left(\frac{l_2^{k''} - l_4^{k''}}{l_2 - l_4}\right) & \sum_{k''=1}^{2}\binom{2}{k''}(-1)^{k''-1}I_{2-k''}\left(\frac{l_3^{k''} - l_4^{k''}}{l_3 - l_4}\right) & \mu^{(2)} + I_2 + \sum_{k''=1}^{2}\binom{2}{k''} I_{k''}\,(-l_4)^{k''} \\ \sum_{k''=1}^{3}\binom{3}{k''}(-1)^{k''-1}I_{3-k''}\left(\frac{l_1^{k''} - l_4^{k''}}{l_1 - l_4}\right) & \sum_{k''=1}^{3}\binom{3}{k''}(-1)^{k''-1}I_{3-k''}\left(\frac{l_2^{k''} - l_4^{k''}}{l_2 - l_4}\right) & \sum_{k''=1}^{3}\binom{3}{k''}(-1)^{k''-1}I_{3-k''}\left(\frac{l_3^{k''} - l_4^{k''}}{l_3 - l_4}\right) & \mu^{(3)} + I_3 + \sum_{k''=1}^{3}\binom{3}{k''} I_{k''}\,(-l_4)^{k''} \\ \sum_{k''=1}^{4}\binom{4}{k''}(-1)^{k''-1}I_{4-k''}\left(\frac{l_1^{k''} - l_4^{k''}}{l_1 - l_4}\right) & \sum_{k''=1}^{4}\binom{4}{k''}(-1)^{k''-1}I_{4-k''}\left(\frac{l_2^{k''} - l_4^{k''}}{l_2 - l_4}\right) & \sum_{k''=1}^{4}\binom{4}{k''}(-1)^{k''-1}I_{4-k''}\left(\frac{l_3^{k''} - l_4^{k''}}{l_3 - l_4}\right) & \mu^{(4)} + I_4 + \sum_{k''=1}^{4}\binom{4}{k''} I_{k''}\,(-l_4)^{k''} \end{pmatrix} . \qquad (6.23a)$$

Next we will carry out the following sequence of elementary row operations on the above matrix:

1) Multiply the first row by $-2I_1$ and add to the second row.
2) Multiply the first row by $-3I_2$ and add to the third row.
3) Multiply the first row by $-4I_3$ and add to the fourth row.
4) Multiply the second row by $-3I_1$ and add to the third row.
5) Multiply the second row by $-6I_2$ and add to the fourth row.
6) Multiply the third row by $-4I_1$ and add to the fourth row.
7) Multiply the third row by l_4 and add to the fourth row.
8) Multiply the second row by l_4 and add to the third row.
9) Multiply the first row by l_4 and add to the second row.

After these elementary row operations, which are equivalent to multiplying the original matrix by a triangular matrix whose determinant is one, our original matrix reduces to

$$\begin{pmatrix} 1 & 1 & 1 & K_1 - l_4 \\ -l_1 & -l_2 & -l_3 & K_2 + K_1\, l_4 \\ l_1^2 & l_2^2 & l_3^2 & K_3 + K_2\, l_4 \\ -l_1^3 & -l_2^3 & -l_3^3 & K_4 + K_3\, l_4 \end{pmatrix} \qquad (6.23b)$$

where K_1, K_2, and K_3 have already been defined, and

$$K_4 = \mu^{(4)} + I_4 - 4I_3K_1 - 6I_2K_2 - 4I_1K_3.$$

Since we want the determinant of $\begin{pmatrix} L' & -V \\ 1'^T & -v \end{pmatrix}$, which is equal to the determinant of (6.23b), we will evaluate this determinant by expanding the determinant of (6.23b) about the elements of the last row giving us, after extensive algebra,

$$(-1)^3\left(\prod_{\substack{k,k'=1 \\ k<k'}}^{3} (1_k - 1_{k'})\right)(K_4 + K_3\Lambda_1^{(C)} + K_2\Lambda_2^{(C)} + K_1\Lambda_3^{(C)} - \Lambda_4^{(C)}),$$

where $C = \{1_1, 1_2, 1_3, 1_4\}$. This last expression is equal to zero by equation (6.20).

By an inductive argument we may show that the general matrix $\begin{pmatrix} L'_T & V \\ 1' & -v \end{pmatrix}$ may similarly be reduced to

$$\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 & K_1 - 1_n \\
-1_1 & -1_2 & -1_3 & \cdots & -1_{n-1} & K_2 + K_1 1_n \\
1_1^2 & 1_2^2 & 1_3^2 & \cdots & 1_{n-1}^2 & K_3 + K_2 1_n \\
-1_1^3 & -1_2^3 & -1_3^3 & \cdots & -1_{n-1}^3 & K_4 + K_3 1_n \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
(-1_1)^{n-1} & (-1_2)^{n-1} & (-1_3)^{n-1} & \cdots & (-1_{n-1})^{n-1} & K_n + K_{n-1} 1_n
\end{pmatrix}, \tag{6.24}$$

and, in addition, an inductive argument may be used to show that the determinant of the above matrix is equal to

$$(-1)^{n-1}\ |D|\ (K_n+K_{n-1}\Lambda_1^{(C)}+K_{n-2}\Lambda_2^{(C)}+K_{n-3}\Lambda_3^{(C)}+\dots$$
$$+K_2\Lambda_{n-2}^{(C)}+K_1\Lambda_{n-1}^{(C)}-\Lambda_n^{(C)})\ , \tag{6.25}$$

which is a generalization of (6.23c). In the above expression $K_1, K_2,\dots,K_{n-1}$, and D have already been defined, $C = \{1_1,1_2,\dots,1_n\} = C^* + \{1_n\}$, and $K_n = \mu^{(n)}+I_n-\binom{n}{1}I_{n-1}K_1-\binom{n}{2}I_{n-2}K_2-\dots -\binom{n}{n-2}I_2K_{n-2}-\binom{n}{n-1}I_1K_{n-1}$. Since $|D| \neq 0$, by using equation (6.20) we obtain for Case 1 the linear combination of the elementary symmetric functions of the elements of C given by:

$$K_{n-1}\Lambda_1^{(C)}+K_{n-2}\Lambda_2^{(C)}+K_{n-3}\Lambda_3^{(C)}+\dots+K_1\Lambda_{n-1}^{(C)}-\Lambda_n^{(C)} = -\ K_n. \tag{6.26a}$$

Next we will consider the regression model designated as Case 2, which may be described by the expression

$$E(Y(z)) = \alpha_1 \exp(-\lambda_1 e^z)+\alpha_2 \exp(-\lambda_2 e^z)+\dots$$
$$+\alpha_{n-1} \exp(-\lambda_{n-1}e^z)-(\alpha_1+\dots+\alpha_{n-1}) \exp(-\lambda_n e^z).$$

It should be noted that the subscript i has still been left off of the regression equation in order to simplify the notation for this particular portion of the discussion. In addition, since most of the conclusions for this case will be obtained by steps that are very similar to those used for Case 1, we will not include as many of the detailed steps here as we did for Case 1. However, we will refer back to the corresponding portions of the discussion for Case 1 so that the reader may be able to add the intermediate steps.

Again we begin our estimation scheme by evaluating the following integral:

$$\mu^{(k')} = \int_{-\infty}^{\infty} z^{k'} dE(Y(z))$$

$$= \int_{-\infty}^{\infty} z^{k'} d\left\{\sum_{k=1}^{n-1} \alpha_k \exp(-\lambda_k e^z) - (\alpha_1+\ldots+\alpha_k) \exp(-\lambda_n e^z)\right\}$$

$$= \int_{-\infty}^{\infty} z^{k'} \left\{ -\sum_{k=1}^{n-1} \alpha_k \lambda_k e^z \exp(-\lambda_k e^z)\right.$$

$$\left. + (\alpha_1+\ldots+\alpha_{n-1}) \lambda_n e^z \exp(-\lambda_n e^z) \right\} dz \tag{6.27}$$

for $k' = 1,2,\ldots,n-1$. Using the same substitution and procedure to evaluate the above set of integrals as we used to derive the set of equations given in (6.9a), we obtain the following set of equations:

$$\sum_{k=1}^{n-1}\left(\sum_{k''=1}^{k'} \binom{k'}{k''} I_{k'-k''} (-1)^{k''-1}((\ln\lambda_k)^{k''} - (\ln\lambda_n)^{k''})\right)\alpha_k$$
$$= \mu^{(k')} \tag{6.28a}$$

for $k' = 1,2,\ldots,n-1$, where I_1, I_2, $I_3,\ldots$ have been defined earlier. Using the same substitution that we used earlier to go from equation (6.9a) to (6.9b), equation (6.28a) becomes:

$$\sum_{k=1}^{n-1}\left(\sum_{k''=1}^{k'} \binom{k'}{k''} I_{k'-k''} (-1)^{k''-1}\left(\frac{l_k^{k''} - l_n^{k''}}{l_k - l_n}\right)\right)\alpha_k' = \mu^{(k')} \tag{6.28b}$$

for $k' = 1,2,\ldots,n-1$.

Using a matrix notation similar to that used to obtain equation (6.9c), we may write equation (6.28b) as:

$$L' \alpha' = V' \tag{6.28c}$$

where L' and α' have been defined earlier, and

$$V' = (\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(n-1)})^T .$$

If we premultiply both sides of equation (6.28c) by the triangular matrix T mentioned in connection with equation (6.12), then equation (6.28c) becomes

$$TL' \alpha' = TV' \text{ or } D \alpha' = TV'$$

where

$$TV' = \begin{pmatrix} K_1' \\ K_2' + K_1' 1_n \\ K_3' + K_2' 1_n \\ \cdot \\ \cdot \\ \cdot \\ K_{n-1}' + K_{n-2}' 1_n \end{pmatrix} \tag{6.29}$$

and

$$K_1' = \mu^{(1)}$$

$$K_2' = \mu^{(2)} - 2I_1K_1'$$

$$K_3' = \mu^{(3)} - 3I_2K_1' - 3I_1K_2'$$

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$K'_{n-1} = \mu^{(n-1)} - \binom{n-1}{1} I_{n-2}K_1' - \binom{n-1}{2} I_{n-3}K_2' - \cdots - \binom{n-1}{n-2} I_1K'_{n-2}.$$

(6.30)

For Case 2 the solution for α' is given by

$$\alpha' = D^{-1} TV' \tag{6.31a}$$

or

$$\alpha'_{k'} = (-1)^{n+k'-1} \frac{|D^{(k')}|}{|D|} \sum_{k=1}^{n-1} \Lambda_{n-k-1}^{(C^*_{(k')})} (K'_k + K'_{k-1} 1_n)$$

$$= (-1)^{n+k'-1} \frac{|D^{(k')}|}{|D|} \sum_{k=1}^{n-1} K'_k \Lambda_{n-k-1}^{(C_{(k')})} \tag{6.31b}$$

for $k' = 1,2,\ldots,n-1$, where $K'_o = 0$, $\Lambda_o^{(C_{(k')})} = 1$ and

$C_{(k')} = C^*_{(k')} + \{1_n\} = \{1_1, 1_2, \ldots, 1_n\} - \{1_{k'}\}$. The next step will be to substitute the expressions for the α'_k 's into the equation

$$\sum_{k=1}^{n-1} \left(\sum_{k''=1}^{n} \binom{n}{k''} I_{n-k''} (-1)^{k''-1} \left(\frac{1_k^{k''} - 1_n^{k''}}{1_k - 1_n} \right) \right) \alpha'_k = \mu^{(n)} . \tag{6.32}$$

It should be noted that equation (6.32) is the same as equation (6.28b) with $k' = n$. With the above substitutions the following relation is satisfied:

$$\begin{vmatrix} L' & -V' \\ 1'^T & -\mu^{(n)} \end{vmatrix} = 0 \tag{6.33a}$$

where L', V', $1'$, and $\mu^{(n)}$ have already been defined. Using the same type of elementary row operations that we used to obtain equation (6.24), which we indicated earlier were equivalent to multiplication by a triangular matrix whose determinant equals one, the matrix $\begin{pmatrix} L' & -V' \\ 1'^T & -\mu^{(n)} \end{pmatrix}$ may be reduced to

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & K_1' \\ -1_1 & -1_2 & -1_3 & \cdots & -1_{n-1} & K_2' + K_1' 1_n \\ 1_1^2 & 1_2^2 & 1_3^2 & \cdots & 1_{n-1}^2 & K_3' + K_2' 1_n \\ -1_1^3 & -1_2^3 & -1_3^3 & \cdots & -1_{n-1}^3 & K_4' + K_3' 1_n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ (-1_1)^{n-2} & (-1_2)^{n-2} & (-1_3)^{n-2} & \cdots & (-1_{n-1})^{n-2} & K_{n-1}' + K_{n-2}' 1_n \\ (-1_1)^{n-1} & (-1_2)^{n-1} & (-1_3)^{n-1} & \cdots & (-1_{n-1})^{n-1} & K_n' + K_{n-1}' 1_n \end{pmatrix}, \tag{6.33b}$$

where K_1', K_2',...,K_{n-1}' have been defined earlier and

$$K_n' = \mu^{(n)} - \binom{n}{1} I_{n-1} K_1' - \binom{n}{2} I_{n-2} K_2' - \cdots - \binom{n}{n-1} I_1 K_{n-1}' .$$

The equation given by (6.33a) is equivalent to the determinant of the matrix given in (6.33b) being set equal to zero. If we expand the determinant of this matrix about the elements of the last row and use an inductive argument similar to that used to obtain equation (6.25), then we find that equation (6.33a) is equivalent to

$$(-1)^{n-1}|D|(K_n' + K_{n-1}'\Lambda_1^{(C)} + K_{n-2}'\Lambda_2^{(C)} + \cdots + K_2'\Lambda_{n-2}^{(C)} + K_1'\Lambda_{n-1}^{(C)}) = 0. \tag{6.33c}$$

Since we are assuming that $|D| \neq 0$, equation (6.33c) is equivalent to

$$K'_{n-1}\Lambda_1^{(C)}+K'_{n-2}\Lambda_2^{(C)}+\ldots+K'_2\Lambda_{n-2}^{(C)}+K'_1\Lambda_{n-1}^{(C)} = -K'_n. \tag{6.33d}$$

At this point we will add the appropriate subscripts to the K's and K''s in equations (6.26a) and (6.33d) so that we will know from which regression equation our linear combination of the elementary symmetric functions arose, and combine the above results into the following theorem:

Theorem 6.1: If a regression model may be specified by equation (6.5c), and if it is assumed that $\sum_{i=1}^{n+1} E(Y_i(z)) = 1$ for all z, then by the evaluation, for each value of $i = 1,2,\ldots,n$, of the set of integrals

$$\mu_i^{(k')} = \int_{-\infty}^{\infty} z^{k'}\, dE(Y_i(z))$$

for $k' = 1,2,\ldots,n$, the following set of linear equations in the elementary symmetric functions of $\ln\lambda_1$, $\ln\lambda_2,\ldots,\ln\lambda_n$ are satisfied:

$$K_{i,n-1}\Lambda_1^{(C)}+K_{i,n-2}\Lambda_2^{(C)}+K_{i,n-3}\Lambda_3^{(C)}+\ldots+K_{i1}\Lambda_{n-1}^{(C)}-\Lambda_n^{(C)} = -K_{in} \tag{6.26b}$$

for $i = 1,2$, and

$$K'_{i,n-1}\Lambda_1^{(C)}+K'_{i,n-2}\Lambda_2^{(C)}+K'_{i,n-3}\Lambda_3^{(C)}+\ldots+K'_{i2}\Lambda_{n-2}^{(C)}+K'_{i1}\Lambda_{n-1}^{(C)} = -K'_{in} \tag{6.33e}$$

for $i = 3,4,\ldots,n$. The K's, K''s, and the set C have all been defined earlier in connection with equations (6.26a) and (6.33d).

From the equations (6.26b) and (6.33e) we have a system of n equations that are linear in the n unknown quantities $\Lambda_1^{(C)}, \Lambda_2^{(C)}, \ldots, \Lambda_n^{(C)}$, i.e. the n elementary symmetric functions of $\ln\lambda_1, \ln\lambda_2, \ldots, \ln\lambda_n$. Therefore we can solve this system of equations for the quantities $\Lambda_1^{(C)}, \Lambda_2^{(C)}, \ldots, \Lambda_n^{(C)}$ in terms of the K's and K´'s. From the solutions for $\Lambda_1^{(C)}, \Lambda_2^{(C)}, \ldots, \Lambda_n^{(C)}$, which will be ratios of linear combinations of the K's and K´'s, we may find the solutions for $\ln\lambda_1, \ln\lambda_2, \ldots, \ln\lambda_n$ by obtaining the n roots of the polynomial

$$w^n - \Lambda_1^{(C)} w^{n-1} + \Lambda_2^{(C)} w^{n-2} - \Lambda_3^{(C)} w^{n-3} + \ldots + (-1)^{n-1}\Lambda_{n-1}^{(C)} w + (-1)^n \Lambda_n^{(C)} = 0. \tag{6.34}$$

Using the relations $\lambda_1 = e^{\ln\lambda_1}, \lambda_2 = e^{\ln\lambda_2}, \ldots, \lambda_n = e^{\ln\lambda_n}$ we may then use the n roots of equation (6.34) to obtain solutions for $\lambda_1, \lambda_2, \ldots, \lambda_n$. We note that the solutions for $\lambda_1, \lambda_2, \ldots, \lambda_n$ are functions of the K's and K´'s, which, in turn, are functions of known constants and the unknown quantities $\mu_i^{(k')}$ for $i, k' = 1,2,\ldots,n$. Hence in order to obtain estimators of the exponential parameters we will propose estimators for the $\mu_i^{(k')}$'s.

On each of the regression equations of our model, the observations will be taken according to the same procedure as described in Section 6.2 for the single exponential model, i.e. the $2M'+1$ or $2M'$ observations on each equation are taken for the values of the independent variable given by $z_j = z_o + jd$ or $z_j = z_o + d(j + \frac{1}{2})$, respectively. Let y_{ij} represent the j^{th} observation on the i^{th} equation. Then by considering the definition of a Riemann-Stieltjes integral the estimator of $\mu_i^{(k')}$ is given by

$$\hat{\mu}_i^{(k')} = \sum_{j=-M'}^{M''-1} \left(\frac{z_j+z_{j+1}}{2}\right)^{k'} \Delta y_{ij}, \tag{6.35a}$$

where $\Delta y_{ij} = y_{i,j+1} - y_{ij}$, z_j has been defined earlier for both situations when an even or odd number of observations have been taken, and $M'' = M'-1$ if an even number of observations are taken but $M'' = M'$ if an odd number of observations are taken. Since we will be interested in investigating some of the asymptotic properties of our estimators of the exponential parameters, we will assume that M' is large enough so that $y_{i,-M'} = E(Y_i(-\infty))$ and $y_{i,M'-1}$ or $y_{i,M'}$ $= E(Y_i(\infty))$ depending upon whether an even or odd number of observations has been taken. With this last assumption the estimator $\hat{\mu}_i^{(k')}$ simplifies to the following expression when an odd number of observations have been taken:

$$\hat{\mu}_i^{(k')} = (z_o - M'd + \tfrac{d}{2})^{k'} E(Y_i(-\infty))$$

$$+ \sum_{j=-M+1}^{M'-2} \left\{ (z_o+jd-\tfrac{d}{2})^{k'} - (z_o+jd+\tfrac{d}{2})^{k'} \right\} y_{ij}$$

$$+ (z_o+M'd-\tfrac{d}{2})^{k'} E(Y_i(+\infty)). \tag{6.35b}$$

If the i^{th} regression equation is of the form specified by Case 1 then the expression for $\hat{\mu}_i^{(k')}$ simplifies to

$$\hat{\mu}_i^{(k')} = \sum_{j=-M'+1}^{M'-1} \left\{ (z_o + d(j - \tfrac{1}{2}))^{k'} - (z_o + d(j + \tfrac{1}{2}))^{k'} \right\} y_{ij}$$

$$+ (z_o + d(M' - \tfrac{1}{2}))^{k'} , \tag{6.35c}$$

and for a Case 2 regression equation this reduces to

$$\hat{\mu}_i^{(k')} = \sum_{j=-M'+1}^{M'-1} \left\{ (z_o + d(j - \tfrac{1}{2}))^{k'} - (z_o + jd + \tfrac{d}{2})^{k'} \right\} y_{ij} . \tag{6.35d}$$

For the situation when an even number of observations have been taken, equations (6.35c) and (6.35d) reduce respectively to

$$\hat{\mu}_i^{(k')} = \sum_{j=-M'+1}^{M'-2} \left\{ (z_o + jd)^{k'} - (z_o + jd + d)^{k'} \right\} y_{ij} + (z_o + d(M'-1))^{k'}$$

and

$$\hat{\mu}_i^{(k')} = \sum_{j=-M'+1}^{M'-2} \left\{ (z_o + jd)^{k'} - (z_o + jd + d)^{k'} \right\} y_{ij} .$$

Hence by using the estimators of $\mu_i^{(k')}$ given by equation (6.35a) in the expressions for the K's and K''s, we may obtain our estimators of the λ_k's for $k = 1, 2, \ldots, n$. The estimators of the linear parameters in our regression model may be found by the same technique as we used in Section 5.3 of Chapter 5.

6.4 Some properties of the generalized Spearman estimators

During the presentation and development of the generalized estimation procedure in Section 6.3, the only assumption that we used concerning the random variables ε_{ij} was the assumption that $E(\varepsilon_{ij})=0$

for all i and j, and the assumption that M' is large enough to that $y_{i,-M'} = E(Y_i(-\infty))$ and $y_{i,M'-1}$ or $y_{i,M'} = E(Y_i(\infty))$. However, before we investigate some of the properties of our estimators we will make some additional assumptions about the random variables ε_{ij} and about the spacing of our observations. These assumptions may be stated as follows:

1) For fixed i, the random variables ε_{ij}, where $i = 1,2,\ldots,n$ and $j = 0, \pm 1, \pm 2,\ldots,\pm M'$ or $j = 0,\pm 1,\pm 2,\ldots,\pm(M'-1), -M'$, are uncorrelated with $E(\varepsilon_{ij}) = 0$ and finite variance such that Var $(\varepsilon_{iM'})$ tends to zero as $M' \longrightarrow \infty$.

2) For $i \neq i'$ and $j \neq j'$, the random variables ε_{ij} and $\varepsilon_{i'j'}$ are uncorrelated.

3) For $k' = 1,2,\ldots,n$ and $i = 1,2,\ldots,n$, the following limit exists:

$$\lim_{M' \longrightarrow \infty} d \sum_{j=-M'+1}^{M''-2} \left[P_{k'-1}(z_j) \right]^2 \mathrm{Var}(\varepsilon_{ij}) \tag{6.36}$$

where M'' has been defined earlier and

$$dP_{k'-1}(z_j) = \left\{ \left(\frac{z_j+z_{j-1}}{2}\right)^{k'} - \left(\frac{z_{j+1}+z_j}{2}\right)^{k'} \right\} = (z_j - \frac{d}{2})^{k'} - (z_j + \frac{d}{2})^{k'}$$

$$= - d \left\{ (z_j - \frac{d}{2})^{k'-1} + (z_j - \frac{d}{2})^{k'-2} (z_j + \frac{d}{2}) + (z_j - \frac{d}{2})^{k'-3} (z_j + \frac{d}{2})^2 + \right.$$

$$\left. \ldots + (z_j - \frac{d}{2})(z_j + \frac{d}{2})^{k'-2} + (z_j + \frac{d}{2})^{k'-1} \right\}. \tag{6.37}$$

4) For all j, as $M' \longrightarrow \infty$ we want $z_{j+1} - z_j \longrightarrow 0$, i.e. $\lim_{M' \longrightarrow \infty} d = 0$.

5) As $M' \longrightarrow \infty$ we want $z_{M''}-z_{-M'} \longrightarrow \infty$, which is equivalent to

$\lim_{M' \longrightarrow \infty} dM' = \infty$.

All of the assumptions of the above lemma are satisfied in some reasonable and practical situations although the third assumption may be difficult to verify. As one particular example where these assumptions are reasonable, we will consider the case when $i = n = k' = 2$ and we will assume that the random variables ε_{2j} are independent with $E(\varepsilon_{2j}) = 0$ and Var $(\varepsilon_{2j}) = \frac{p_{2j}(1-p_{2j})}{n*}$ where $p_{2j} = \alpha_{21}\exp(-\lambda_1 e^{z_j}) + (1-\alpha_{21})\exp(-\lambda_2 e^{z_j})$. This particular case might arise when we use the normal approximation to the distribution of a binomially distributed random variable. For this situation

$$P_1(z_j) = (z_j - \frac{d}{2}) + (z_j + \frac{d}{2}) = 2z_j$$

and equation (6.36) becomes

$$\lim_{M' \longrightarrow \infty} d \sum_{j=-M'+1}^{M''-2} (2z_j)^2 \frac{p_{2j}(1-p_{2j})}{n*}$$

$$= \frac{4}{n*} \int_{-\infty}^{\infty} z^2(\alpha_{21}\exp(-\lambda_1 e^z)+(1-\alpha_{21})\exp(-\lambda_2 e^z))$$

$$(1-\alpha_{21}\exp(-\lambda_1 e^z) - (1-\alpha_{21})\exp(-\lambda_2 e^z))\, dz.$$

By making the substitution $t = \exp(-e^z)$ and after extensive algebra we can show that the above integral is equal to

$$\frac{4}{n^*}\left[\alpha_{21}\left\{(\ln\lambda_2-\ln\lambda_1)\left[I_2+\gamma(\ln\lambda_2+\ln\lambda_1)+\frac{1}{3}((\ln\lambda_2)^2+\ln\lambda_1\ln\lambda_2+(\ln\lambda_1)^2)\right]\right.\right.$$

$$+\ 2(\ln(\lambda_1+\lambda_2)-\ln\lambda_2-\ln 2)\left[I_2+\gamma(\ln(\lambda_1+\lambda_2)+\ln\lambda_2+\ln 2)\right.$$

$$\left.\left.+\ \frac{1}{3}\Big((\ln(\lambda_1+\lambda_2))^2+\ln(\lambda_1+\lambda_2)\ln\lambda_2+\ln(\lambda_1+\lambda_2)\ln 2+(\ln\lambda_2)^2+(\ln 2)^2+2\ln 2\ln\lambda_2\Big)\right]\right\}$$

$$+\ \ln\lambda_2\left[I_2+\gamma(2\ln\lambda_2+\ln 2)+\frac{1}{3}((\ln 2)^2+3\ln 2\ln\lambda_2+3(\ln\lambda_2)^2)\right]$$

$$+\ \alpha_{21}^2\left\{(\ln 2+\ln\lambda_1-\ln(\lambda_1+\lambda_2))\left[I_2+\gamma(\ln 2+\ln\lambda_1+\ln(\lambda_1+\lambda_2))\right.\right.$$

$$\left.+\ \frac{1}{3}\Big((\ln 2)^2+2\ln 2\ln\lambda_1+(\ln\lambda_1)^2+\ln 2\ln(\lambda_1+\lambda_2)+\ln\lambda_1\ln(\lambda_1+\lambda_2)+(\ln(\lambda_1+\lambda_2))^2\Big)\right]$$

$$+\ (\ln 2+\ln\lambda_2-\ln(\lambda_1+\lambda_2))\left[I_2+\gamma(\ln 2+\ln\lambda_2+\ln(\lambda_1+\lambda_2))\right.$$

$$\left.\left.\left.+\ \frac{1}{3}\Big((\ln 2)^2+2\ln 2\ln\lambda_2+(\ln\lambda_2)^2+\ln 2\ln(\lambda_1+\lambda_2)+\ln\lambda_2\ln(\lambda_1+\lambda_2)+(\ln(\lambda_1+\lambda_2))^2\Big)\right]\right\}\right]$$

where I_2 and γ have been defined earlier. Since α_{21}, λ_1, and λ_2 are nonzero constants, it is obvious that the above expression is finite. Hence we see that the third assumption in a practical example holds and can be verified.

At this point we want to show that the estimators that have been proposed in Section 6.3 for the exponential parameters are consistent estimators under the five assumptions given above. In order to prove this we will need the following lemma:

Lemma 6.2: If $\hat{\mu}_i^{(k')}$, as given by equation (6.35a), is the proposed estimator of $\mu_i^{(k')}$ for i, $k'=1,2,\ldots,n$, then $\hat{\mu}_i^{(k')}-\mu_i^{(k')}$ tends in probability to zero under the assumptions stated at the beginning of Section 6.4.

Proof: From the specification of our regression model by equations (6.5a) through (6.5c) we may write

$$\hat{\mu}_i^{(k')}-\mu_i^{(k')}=\sum_{j=-M'}^{M''-1}\left(\frac{z_j+z_{j+1}}{2}\right)^{k'}\Delta E(Y_i(z_j))-\mu_i^{(k')}+\sum_{j=-M'}^{M''-1}\left(\frac{z_j+z_{j+1}}{2}\right)^{k'}\Delta\varepsilon_{ij}. \tag{6.38}$$

Using the definition of the Riemann-Stieltjes integral (see Olmstead [1959], page 179) and the fourth and fifth assumptions of this lemma, we can write

$$\lim_{M'\longrightarrow\infty}\left\{\sum_{j=-M'}^{M''-1}\left(\frac{z_j+z_{j+1}}{2}\right)^{k'}\Delta E(Y_i(z_j)) - \mu_i^{(k')}\right\} = 0. \tag{6.39}$$

Next consider the term

$$\sum_{j=-M'}^{M''-1}\left(\frac{z_j+z_{j+1}}{2}\right)^{k'}\Delta\varepsilon_{ij}. \tag{6.40}$$

Since we have developed this estimation procedure under the assumption that M' is large enough so that $y_{i,-M'}=E(Y_i(-\infty))$ and $y_{i,M''}=E(Y_i(+\infty))$, where $M''=M'$ if an odd number of observations has been taken but $M''=M'-1$ if an even number has been taken, we may write equation (6.40) as

$$\sum_{j=-M'+1}^{M''-2}\left\{\left(\frac{z_j+z_{j-1}}{2}\right)^{k'}-\left(\frac{z_{j+1}+z_j}{2}\right)^{k'}\right\}\varepsilon_{ij}. \tag{6.41}$$

If we can show that the variance of (6.41) converges to zero as $M'\longrightarrow\infty$, then we may apply a form of Tchebycheff's theorem given by

Cramér ([1946], page 253) to conclude that this expression converges in probability to zero. Using the first and second assumptions of this lemma, the variance of (6.41) is given by

$$d^2 \sum_{j=-M'+1}^{M''-2} \left\{P_{k'-1}(z_j)\right\}^2 \text{Var}\,(\varepsilon_{ij}) \tag{6.42}$$

where $P_{k'-1}(z_j)$ has been defined in equation (6.37). From the third and fourth assumptions of this lemma given at the beginning of Section 6.4, the expression in equation (6.42) tends to zero as $M' \longrightarrow \infty$, since $d \longrightarrow 0$ and the limit given by equation (6.36) exists. Hence by combining the above results we may conclude that $\hat{\mu}_i^{(k')} - \mu_i^{(k')}$ converges in probability to zero and $\hat{\mu}_i^{(k')}$ is a consistent estimator of $\mu_i^{(k')}$.

The lemma, which is presented next, is an interesting result concerning the fourth and fifth assumptions used in Lemma 6.2.

Lemma 6.3: The condition $d=0(M'^{-\varepsilon})$ for $0 < \varepsilon < 1$ is a sufficient condition for the fourth and fifth assumptions used in Lemma 6.2 to hold.

Proof: From the hypothesis of our lemma, we know that $d=0(M'^{-\varepsilon})$. Therefore d is at most of the order of $M'^{-\varepsilon}$. However $\lim_{M' \longrightarrow \infty} M'^{-\varepsilon}=0$ for $0 < \varepsilon < 1$. Hence $\lim_{M' \longrightarrow \infty} d=0$. Also since $d=0(M'^{-\varepsilon})$, we know that $\lim_{M' \longrightarrow \infty} \frac{d}{M'^{-\varepsilon}} = \lim_{M' \longrightarrow \infty} dM'^{\varepsilon} = \text{constant} < \infty$. Therefore

$$\lim_{M' \longrightarrow \infty} dM' = \lim_{M' \longrightarrow \infty} dM'^{\varepsilon} M'^{1-\varepsilon} = \infty.$$

Hence from our hypothesis we note that the fourth and fifth assumptions of Lemma 6.2 are satisfied.

From the estimation equations for the elementary symmetric functions, $\Lambda_r^{(C)}$, $r=1,2,\ldots,n$, of $\ln\lambda_1$, $\ln\lambda_2,\ldots,\ln\lambda_n$ given in Section 6.3, we can easily see that these functions are rational functions of the $\hat{\mu}_i^{(k')}$'s. Hence from Slutsky's theorem (see Cramér [1946], page 254) we conclude that $L_r^{(C)}$, the estimator of $\Lambda_r^{(C)}$ for $r = 1,2,\ldots,n$, is a consistent estimator of $\Lambda_r^{(C)}$. From the method that we used in Section 6.3 to solve for the parameters $\lambda_1,\lambda_2,\ldots,\lambda_n$, these parameters will be continuous functions of the $\Lambda_r^{(C)}$ for $r=1, 2,\ldots,n$. Since the estimators $\hat{\lambda}_1,\hat{\lambda}_2,\ldots,\hat{\lambda}_n$ of the parameters $\lambda_1,\lambda_2,\ldots,\lambda_n$ are found by substituting the expressions for $L_r^{(C)}$ in place of $\Lambda_r^{(C)}$, these estimators will be continuous functions of the $L_r^{(C)}$ for $r=1,2,\ldots,n$, and we may conclude that the estimators $\hat{\lambda}_1,\hat{\lambda}_2,\ldots,\hat{\lambda}_n$ are consistent estimators of $\lambda_1,\lambda_2,\ldots,\lambda_n$ (see Wilks [1962], page 103). From the above discussion we have the following theorem:

Theorem 6.4: Under the assumptions of Lemma 6.2, the estimators $\hat{\lambda}_1,\hat{\lambda}_2,\ldots,\hat{\lambda}_n$, which have been found by the generalized Spearman estimation procedure, of the parameters $\lambda_1,\lambda_2,\ldots,\lambda_n$ in the regression model specified by equations (6.5a) through (6.5c), are consistent.

Since we will be using the same procedure here to estimate the linear parameters of our regression model as we used in Chapter 5, we will also have a theorem corresponding to Theorem 5.4 given by:

Theorem 6.5: The estimators of the linear parameters in the regression model, specified by equations (6.52 a) through (6.5c), that have been found by the generalized Spearman estimation procedure of Section 6.3 are consistent, under the following assumptions:

1) The assumptions of Lemma 6.2 are satisfied.

2) The random variables ε_{ij} are normally distributed.

3) The estimators of the linear parameters are continuous functions. Since the proof of this theorem follows closely the proof of Theorem 5.4, we will not repeat it here.

Now that we have established the consistency of the estimators of the parameters in our regression model, we want to determine the asymptotic distribution of the n × 1 vector of estimators of the exponential parameters $\lambda_1,\lambda_2,\ldots,\lambda_n$. Before deriving this asymptotic distribution we will consider the following vector:

$$\hat{\mu}_*^{(*)} - \mu_*^{(*)} = (\hat{\mu}_1^{(1)} - \mu_1^{(1)},\ldots,\hat{\mu}_1^{(n)} - \mu_1^{(n)},\ \hat{\mu}_2^{(1)} - \mu_2^{(1)},\ldots,$$

$$\hat{\mu}_2^{(n)} - \mu_2^{(n)},\ldots,\hat{\mu}_n^{(1)} - \mu_n^{(1)},\ldots,\hat{\mu}_n^{(n)} - \mu_n^{(n)})^T, \tag{6.43}$$

where $\hat{\mu}_i^{(k')}$ and $\mu_i^{(k')}$ have been defined earlier in equation (6.35a) and Theorem 6.1, respectively. Now let us consider

$$(\hat{\mu}_*^{(*)} - \mu_*^{(*)}) \;/\; d^{\frac{3}{2}} \sqrt{M' + M'' - 2} \tag{6.44}$$

where d, M', and M'' have already been defined. Using Lemma 6.3 we can show that $\lim_{M' \to \infty} d^{\frac{3}{2}} \sqrt{M'+M''-2}$ is a constant. Therefore, using Lemma 6.2 and a convergence theorem from Cramér ([1946], page 254), we conclude that the limiting distribution of the expression in equation (6.44) is the same as the limiting distribution of $\varepsilon_{*.}^{(d)} / d^{\frac{3}{2}} \sqrt{M'+M''-2}$, since

$$\frac{\hat{\mu}_{*}^{(*)} - \mu_{*}^{(*)}}{d^{\frac{3}{2}} \sqrt{M'+M''-2}} = \frac{E(\hat{\mu}_{*}^{(*)}) - \mu_{*}^{(*)}}{d^{\frac{3}{2}} \sqrt{M'+M''-2}} + \frac{\varepsilon_{*.}^{(d)}}{d^{\frac{3}{2}} \sqrt{M'+M''-2}}, \tag{6.45}$$

where $\varepsilon_{*.}^{(d)} = \sum_{j=-M'+1}^{M''-2} \varepsilon_{*j}^{(d)}$; $\varepsilon_{*j}^{(d)} = \left(\varepsilon_{1j}^{(d)^T}, \varepsilon_{2j}^{(d)^T}, \ldots, \varepsilon_{nj}^{(d)^T}\right)^T$; and

$\varepsilon_{ij}^{(d)} = -d\varepsilon_{ij}(P_o(z_j), P_1(z_j), \ldots, P_{n-1}(z_j))^T$ where $P_{k'-1}(z_j)$ has been defined earlier. With the above definitions and a direct application of a multivariate form of the central limit theorem stated by Cramér ([1936], pages 113-114), we may prove the following lemma:

Lemma 6.6: Under the assumptions used in Lemma 6.2, the vector $\varepsilon_{*.}^{(d)} / d^{\frac{3}{2}} \sqrt{M'+M''-2}$ has a limiting multivariate normal distribution with mean vector 0 and covariance matrix $\lim_{M' \to \infty} \frac{E(\varepsilon_{*.}^{(d)} \varepsilon_{*.}^{(d)^T})}{d^3(M'+M''-2)}$, provided

1) $$\lim_{M' \to \infty} \frac{E(\varepsilon_{*.}^{(d)} \varepsilon_{*.}^{(d)^T})}{d^3(M'+M''-2)} \qquad (6.46)$$

exists (i.e. the limit of each element of this matrix exists) and not every element of the matrix is equal to zero;

2) $$\lim_{M' \to \infty} \frac{1}{M'+M''-2} \sum_{j=-M'+1}^{M''-2} \int_{(t_1^2+\ldots+t_n^2)>\xi(M'+M''-2)} (t_1^2+t_2^2+\ldots+t_n^2)\, dF_j = 0 \qquad (6.47)$$

for every $\xi > 0$ where F_j is the distribution function of $\varepsilon_{*j}^{(d)}/ d^{\frac{3}{2}}$.

In some practical situations it may be difficult to determine whether the two additional assumptions of Lemma 6.6 are satisfied or not. However, these additional assumptions are satisfied in some reasonable and practical situations, as we will show by considering a particular example. The particular example that we will consider involves the case when

$\varepsilon_{*j}^{(d)}/ d^{\frac{3}{2}} = - \varepsilon_{1j}/ d^{\frac{1}{2}}$ has a N $(0,\sigma_j^2/ d)$ distribution, and this example could arise when a normal approximation is used for the distribution of binomially distributed random variables. This is similar to the example considered earlier in our discussion on the reasonableness of the assumptions stated at the beginning of Section 6.4. The first of these two additional assumptions may be verified in a manner very similar to that used to verify the third assumption given at the beginning of Section 6.4. In fact, we must use that assumption in order to verify whether this first additional assumption is satisfied, since the diagonal elements of (6.46) involve terms of the form given by (6.36). Concerning the second additional assumption, for this particular example equation (6.47) becomes

$$\lim_{M' \to \infty} \frac{1}{M'+M''-2} \sum_{j=-M'+1}^{M''-2} \int_{t_1^2>\xi^2(M'+M''-2)} \frac{t_1^2 e^{-\{t_1^2/(2\sigma_j^2/d)\}}}{\sqrt{2\pi\sigma_j^2/d}} \, dt_1 \, .$$

Let $t_1^* = \dfrac{t_1}{\sqrt{M'+M''-2}}$. Then the above expression can be shown to be less than or equal to

$$\lim_{M' \to \infty} \sum_{j=-M\;+1}^{M''-2} \sigma_j^2/\{d(M'+M''-2)\}.$$

From the assumptions of Lemma 6.2 stated at the beginning of Section 6.4, we conclude that this limit equals zero. Hence we can see that the two additional assumptions of Lemma 6.6 are satisfied in at least one reasonable situation.

Using the conclusion of Lemma 6.6 along with the fact that $(\hat{\mu}_*^{(*)} - \mu_*^{(*)}) \,/\, d^{\frac{3}{2}} (M'+M''-2)^{\frac{1}{2}}$ has the same limiting distribution as $\varepsilon_{*.}^{(d)} / \, d^{\frac{3}{2}} (M'+M''-2)^{\frac{1}{2}}$, we will expand the estimators of the exponential parameters in a Taylor's series about the point $\mu_*^{(*)}$ within the neighborhood $|\hat{\mu}_i^{(k')} - \mu_i^{(k')}| \leq \delta$ for all i and k′ and $\delta > 0$ where the vector $\hat{\mu}_*^{(*)} - \mu_*^{(*)}$ has already been defined earlier.

By doing this we have

$$\hat{\lambda} = \hat{\lambda} \Big|_{\hat{\mu}_*^{(*)} = \mu_*^{(*)}} + F'(\hat{\mu}_*^{(*)} - \mu_*^{(*)}) + G' \tag{6.48}$$

where

$$F' = \begin{pmatrix} F_1' \\ F_2' \\ \cdot \\ \cdot \\ \cdot \\ F_n' \end{pmatrix}; \quad F_k' = \left(\frac{\partial\hat{\lambda}_k}{\partial\hat{\mu}_1^{(1)}}\Big|_{\hat{\mu}_*^{(*)}=\mu_*^{(*)}}, \ldots, \frac{\partial\hat{\lambda}_k}{\partial\hat{\mu}_1^{(n)}}\Big|_{\hat{\mu}_*^{(*)}=\mu_*^{(*)}}, \ldots, \frac{\partial\hat{\lambda}_k}{\partial\hat{\mu}_n^{(1)}}\Big|_{\hat{\mu}_*^{(*)}=\mu_*^{(*)}}, \ldots, \frac{\partial\hat{\lambda}_k}{\partial\hat{\mu}_n^{(n)}}\Big|_{\hat{\mu}_*^{(*)}=\mu_*^{(*)}} \right)$$

for k = 1,2,...,n; and $G' = \begin{pmatrix} G_1' \\ G_2' \\ \cdot \\ \cdot \\ \cdot \\ G_n' \end{pmatrix}$ where

$$G_k' = \frac{1}{2} \left(\sum_{i,k',i',k''} (\hat{\mu}_i^{(k')} - \mu_i^{(k')})(\hat{\mu}_{i'}^{(k'')} - \mu_{i'}^{(k'')}) \frac{\partial^2\hat{\lambda}_k}{\partial\hat{\mu}_i^{(k')}\partial\hat{\mu}_{i'}^{(k'')}} \Big|_{\hat{\mu}_*^{(*)}=\mu_*^{(*)}+\theta_k(\hat{\mu}_*^{(*)}-\mu_*^{(*)})} \right)$$

for k = 1,2,...,n and $|\theta_k| < 1$. Under the assumption that the estimators $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n$ have continuous second order derivatives within the region $|\hat{\mu}_i^{(k')} - \mu_i^{(k')}| \leq \delta$, we may use a proof similar to that used to prove Theorem 5.5 to conclude that

$(\hat{\lambda}-\lambda)/(d^{\frac{3}{2}}\sqrt{M'+M''-2})$ has the same limiting distribution as

$F'(\hat{\mu}_*^{(*)}-\mu_*^{(*)})/(d^{\frac{3}{2}}(M'+M''-2)^{\frac{1}{2}})$. From Lemmas 6.2 and 6.6 we

conclude that the limiting distribution of $(\hat{\lambda}-\lambda)/(d^{\frac{3}{2}}(M'+M''-2)^{\frac{1}{2}})$

is a multivariate normal distribution with mean vector 0 and

covariance matrix

$$\lim_{M' \longrightarrow \infty} \left(\frac{1}{d^3(M'+M''-2)} \right) F'E\left(\varepsilon_{*.}^{(d)} \varepsilon_{*.}^{(d)^T} \right) F'^T . \qquad (6.49)$$

Combining these results we have:

Theorem 6.7: If $\hat{\lambda}$ represents the estimator of the vector λ of exponential parameters of our regression model found by the generalized Spearman estimation procedure presented in Section 6.3, if the elements of $\hat{\lambda}$ have continuous second order derivatives of every kind with respect to the elements of $\hat{\mu}_*^{(*)}$, and if the assumptions of Lemmas 6.2, 6.6 and Theorem 6.4 are met, then $(\hat{\lambda}-\lambda)/(d^{\frac{3}{2}}(M'+M''-2)^{\frac{1}{2}})$ has a limiting multivariate normal distribution with mean vector 0 and covariance matrix given by equation (6.49) as $M' \longrightarrow \infty$.

As we did in Chapter 5, we now propose to obtain an expression for the asymptotic efficiency of the generalized Spearman estimators of the exponential parameters in our regression model. From the ideas presented by Kendall and Stuart [1961], we take as our measure of the asymptotic efficiency of our estimators, the following ratio:

$$\nu = \lim_{M' \longrightarrow \infty} \left\{ \left| E\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T \right| \left| \Omega \right| \right\}^{-1} \tag{6.50}$$

where L represents the likelihood function whose form will be specified and Ω represents the asymptotic covariance matrix of our estimators as obtained from the results of Theorem 6.7.

We will derive the expression for ν when the assumptions concerning the random variables of our regression model given at the beginning of Section 6.4 are satisfied and the likelihood function is specified by

$$L = \prod_{j=-M'}^{M''-1} \frac{1}{(2\pi)^{M'+M''} |\Sigma_j|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(Y_{*j}-E(Y_{*j}))^T \Sigma_j^{-1} (Y_{*j}-E(Y_{*j}))\right\} \tag{6.51}$$

where

$$Y_{*j} = \begin{pmatrix} Y_{1j} \\ Y_{2j} \\ \cdot \\ \cdot \\ \cdot \\ Y_{nj} \end{pmatrix} ; \quad \Sigma_j = E(Y_{*j}-E(Y_{*j}))(Y_{*j}-E(Y_{*j}))^T;$$ and in accordance

with equation (6.5c) and the assumption $\sum_{i=1}^{n+1} E(Y_i(z)) = 1$ for all z,

$$E(Y_{*j}) = \begin{pmatrix} E(Y_{1j}) \\ E(Y_{2j}) \\ \cdot \\ \cdot \\ \cdot \\ E(Y_{nj}) \end{pmatrix}$$

for $j = -M', -M'+1, \ldots, 0, 1, \ldots, M'-1$, and/or M'. That is, we are assuming that the vectors Y_{*j} are independent each with a multivariate $N(E(Y_{*j}), \Sigma_j)$ distribution. In order to find ν, we need:

$$\ln L = -\sum_{j=-M'}^{M''-1} \ln (2\pi)^{\frac{M'+M''}{2}} - \frac{1}{2} \sum_{j=-M'}^{M''-1} \ln|\Sigma_j|$$

$$- \frac{1}{2} \sum_{j=-M'}^{M''-1} (Y_{*j}-E(Y_{*j}))^T \Sigma_j^{-1} (Y_{*j}-E(Y_{*j})) \tag{6.52}$$

and

$$\frac{\partial \ln L}{\partial \lambda} = \sum_{j=-M'}^{M''-1} D_j \Sigma_j^{-1} (Y_{*j}-E(Y_{*j})) \tag{6.53}$$

where

$$D_j = \begin{pmatrix} \frac{\partial E(Y_{1j})}{\partial \lambda_1} & \cdots & \frac{\partial E(Y_{nj})}{\partial \lambda_1} \\ \vdots & & \vdots \\ \frac{\partial E(Y_{1j})}{\partial \lambda_n} & \cdots & \frac{\partial E(Y_{nj})}{\partial \lambda_n} \end{pmatrix} . \tag{6.54}$$

Hence from the assumptions that we have made at the beginning of Section 6.4 we have:

$$E\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T = \sum_{j=-M'}^{M''-1} D_j \Sigma_j^{-1} D_j^T. \tag{6.55}$$

From Theorem 6.7 we have the asymptotic covariance matrix of $\hat{\lambda}$ given by

$$F' E(\varepsilon_{*.}^{(d)} \varepsilon_{*.}^{(d)T})\, F'^T = \sum_{j=-M'}^{M''-1} F' \,(W_j \otimes \Sigma_j)\, F'^T \tag{6.56}$$

where

$$W_j = d^2 \begin{pmatrix} P_o^2(z_j) & P_o(z_j)P_1(z_j) & \cdots & P_o(z_j)P_{n-1}(z_j) \\ P_o(z_j)P_1(z_j) & P_1^2(z_j) & \cdots & P_1(z_j)P_{n-1}(z_j) \\ \vdots & \vdots & & \vdots \\ P_o(z_j)P_{n-1}(z_j) & P_1(z_j)P_{n-1}(z_j) & \cdots & P_{n-1}^2(z_j) \end{pmatrix}$$

and (6.57)

$$\Sigma_j = \begin{pmatrix} E(\varepsilon_{1j}^2) & E(\varepsilon_{1j}\varepsilon_{2j}) & \cdots & E(\varepsilon_{1j}\varepsilon_{nj}) \\ E(\varepsilon_{1j}\varepsilon_{2j}) & E(\varepsilon_{2j}^2) & \cdots & E(\varepsilon_{2j}\varepsilon_{nj}) \\ \vdots & \vdots & & \vdots \\ E(\varepsilon_{1j}\varepsilon_{nj}) & E(\varepsilon_{2j}\varepsilon_{nj}) & \cdots & E(\varepsilon_{nj}^2) \end{pmatrix} \tag{6.58}$$

and F' has been defined earlier. Therefore we may combine the above results into the following theorem:

Theorem 6.8: When the likelihood function of the vectors Y_{*j} is given by equation (6.51) then the expression for the asymptotic efficiency, defined in (6.50), of the vector of estimators of the exponential parameters reduces to the following expression:

$$\nu = \lim_{M' \to \infty} \left\{ \left| \sum_{j=-M'}^{M''-1} D_j \Sigma_j^{-1} D_j^T \right| \left| \sum_{j=-M'}^{M''-1} F'(W_j \otimes \Sigma_j) F'^T \right| \right\}^{-1} \tag{6.59}$$

In Chapter 7, where a comparison of the various estimation techniques discussed in this research will be made, the expression given by equation (6.59) will be evaluated for some particular regression models.

VII. COMPARISONS AND ILLUSTRATIONS OF THE GENERALIZED ESTIMATION PROCEDURES

7.1 Introduction

In this chapter we will apply the generalized estimation procedures developed in this research to various sets of data concerned with experimental situations. The particular models that we will be considering in this chapter are given by equation (1.1) with $n = m = 2$, where we assume that α_{io} is known. We will work out some of the details of these estimation procedures and in Sections 7.2.2 and 7.2.3, where the generalized partial totals and generalized Spearman estimation techniques will be applied to particular sets of data, we will also determine the generalized least squares estimates of the parameters of interest. This will give us some comparisons of the various techniques for particular sets of data. Finally we will evaluate the expressions for the asymptotic efficiency of our generalized estimation procedures for some special cases, and this will be used as another criterion for comparing these various estimation techniques.

7.2 Application of the estimation procedures to specific examples

7.2.1 Generalized least squares

In this section we apply the generalized least squares procedure to the experimental example presented in the article by Galambos and Cornell [1962] where a mathematical model is formulated to describe sulfate metabolism in patients. We may think of this problem as a three-compartment mammillary (or catenary) system where the radioactive tracer is initially introduced into the first compartment and ultimately accumulates in the second compartment. Figure 7.1 gives a schematic representation of the compartmental system of interest. The

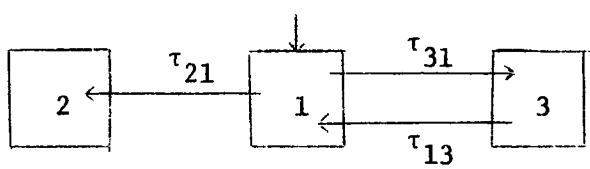

Fig. 7.1--Compartmental model for the data given in Table 7.1

numbered boxes in this figure represent the three compartments, and the τ's in this figure represent the turnover rates or transition probabilities defined in Section 3.1 of Chapter 3. The observable random variables represent the proportion of injected radioactive tracer present in the respective compartments; and since the sum of the expected values of these random variables is always equal to one, there are only two independent regression equations. The set of independent regression equations for this particular example is given by

$$Y_{1j} = \theta_1 e^{-\theta_2 x_j} + (1-\theta_1)e^{-\theta_3 x_j} + \varepsilon_{1j} = f_1(\theta,x_j) + \varepsilon_{1j}$$

$$Y_{2j} = 1-(\theta_1+\theta_4)e^{-\theta_2 x_j} - (1-\theta_1-\theta_4)e^{-\theta_3 x_j} + \varepsilon_{2j}$$

$$= f_2(\theta,x_j) + \varepsilon_{2j} , \qquad (7.1)$$

where $\theta_4 = (\theta_3-\theta_2)\theta_1(1-\theta_1)/[(\theta_3-\theta_2)\theta_1+\theta_2]$, so there are actually only three independent parameters to be estimated. Using the notation of Chapter 4 we have $\theta = (\theta_1,\theta_2,\theta_3)^T$. The data for this particular example is given in Table 7.1, where the values taken on by the random variables Y_{1j} and Y_{2j} are denoted in the columns headed by y_{1j} and y_{2j}, respectively. All of the values given in Table 7.1 were taken from the article by Galambos and Cornell [1962], except the value of y_{1j} for $j = 1$, which was extrapolated from the earlier observations so that we could display the generalized least squares procedure in its simplest form.

The first step in our generalized estimation procedure is to obtain estimates of the quantities $\sigma_{11} = E(\varepsilon_{1j}^2)$, $\sigma_{22} = E(\varepsilon_{2j}^2)$, and $\sigma_{12} = E(\varepsilon_{1j}\varepsilon_{2j})$ for all j. Using the observed proportions, plots are made on semi-logarithmic paper in order to compute the vector of preliminary estimates of the vector θ given by ${}_o\hat{\theta} = (0.381, 0.021, 0.197)^T$. Using the proportions y_{1j} and the vector of preliminary estimates ${}_o\hat{\theta}$, we apply Hartley's modified

TABLE 7.1--Data to be fitted by the generalized least squares procedure

j	x_j	y_{1j}	y_{2j}
1	0.33	0.92	0.03
2	2	0.84	0.10
3	3	0.79	0.14
4	5	0.64	0.21
5	8	0.55	0.30
6	12	0.44	0.40
7	24	0.27	0.54
8	48	0.12	0.66
9	72	0.06	0.71

Gauss-Newton procedure as discussed in Section 4.2 to the first regression equation of (7.1). After applying this technique to the first regression equation of our model, we find the following vector of estimators of the elements of θ: $\hat{\theta}^{(1)} = (0.5752, 0.0322, 0.1816)^T$. From equation (4.14) we recall that $\hat{\sigma}_{11} = \frac{1}{9}\hat{\varepsilon}_{1*}^T\hat{\varepsilon}_{1*}$, where the elements $\hat{\varepsilon}_{1j}$ of the vector $\hat{\varepsilon}_{1*}$ are given in column (2) of Table 7.2. Similarly, using the observations y_{2j} from the second regression equation of our model given in equation (7.1), we find $\hat{\theta}^{(2)} = (0.282, 0.021, 0.195)^T$. The elements $\hat{\varepsilon}_{2j}$ of the vector $\hat{\varepsilon}_{2*}$ defined in Chapter 4 are given

TABLE 7.2--Deviations to be used to estimate covariance matrix

(1) j	(2) $\hat{\varepsilon}_{1j}$	(3) $\hat{\varepsilon}_{2j}$
1	-0.04920	0.01287
2	0.00528	0.00340
3	0.02144	0.00107
4	-0.02093	-0.00339
5	0.00615	-0.00380
6	0.00121	0.00588
7	-0.00087	-0.00485
8	-0.00255	0.00442
9	0.00348	-0.00200

in column (3) of Table 7.2. The sums of squares of the entries in columns (2) and (3) of Table 7.2 after division by N = 9 yield $\hat{\sigma}_{11} = 3.7832 \times 10^{-4}$ and $\hat{\sigma}_{22} = 3.1787 \times 10^{-5}$, respectively. The sum of the cross products of these columns after division by N = 9 gives $\hat{\sigma}_{12} = -6.1292 \times 10^{-5}$.

Having estimated the elements of the symmetric matrix σ_{**}, the estimated matrix is denoted by $\hat{\sigma}_{**}$ and an estimate of Ω is given by $\hat{\Omega} = I \otimes \hat{\sigma}_{**}$. Using the vector ${}_{o}\hat{\theta}$ of preliminary estimates of the vector θ, we determine the elements of the matrices ${}_o f'$ and $y - {}_o f'$ defined in Section 4.3. The regression equations in (7.1) with ${}_o\hat{\theta}$ substituted for θ are used to evaluate $y - {}_o f$, and in order to evaluate the matrix ${}_o f'$ for this example, the derivatives $\partial f_1(\theta, x_j)/\partial\theta_k$ and $\partial f_2(\theta, x_j)/\partial\theta_k$ for $k = 1,2,3$ are calculated with ${}_o\hat{\theta}$ substituted for θ using the derivatives $\partial\theta_4/\partial\theta_k$ for $k = 1,2,3$. Substitution of these quantities into equation (4.13) along with the substitution of $\hat{\Omega}$ for Ω, leads after one iteration to a new vector of estimates ${}_1\hat{\theta} = (0.30558, 0.01870, 0.14350)^T$. Using Hartley's modified procedure, after nine iterations this iterative process gives us $\hat{\theta} = (0.07397, 0.00752, 0.09228)^T$ correct to four decimal places. Graphs showing the original data (x) and the fitted equations are shown in Figures 7.2 and 7.3 for the first and second regression equations, respectively, of (7.1).

7.2.2 Generalized partial totals

In this section we apply the generalized partial totals estimation procedure developed in Chapter 5 to the set of data given in Table 7.3. The data given in this table was manufactured for the regression model given in equation (7.2) below by adding random normal deviates to calculated expected values. The regression model of interest is given as follows:

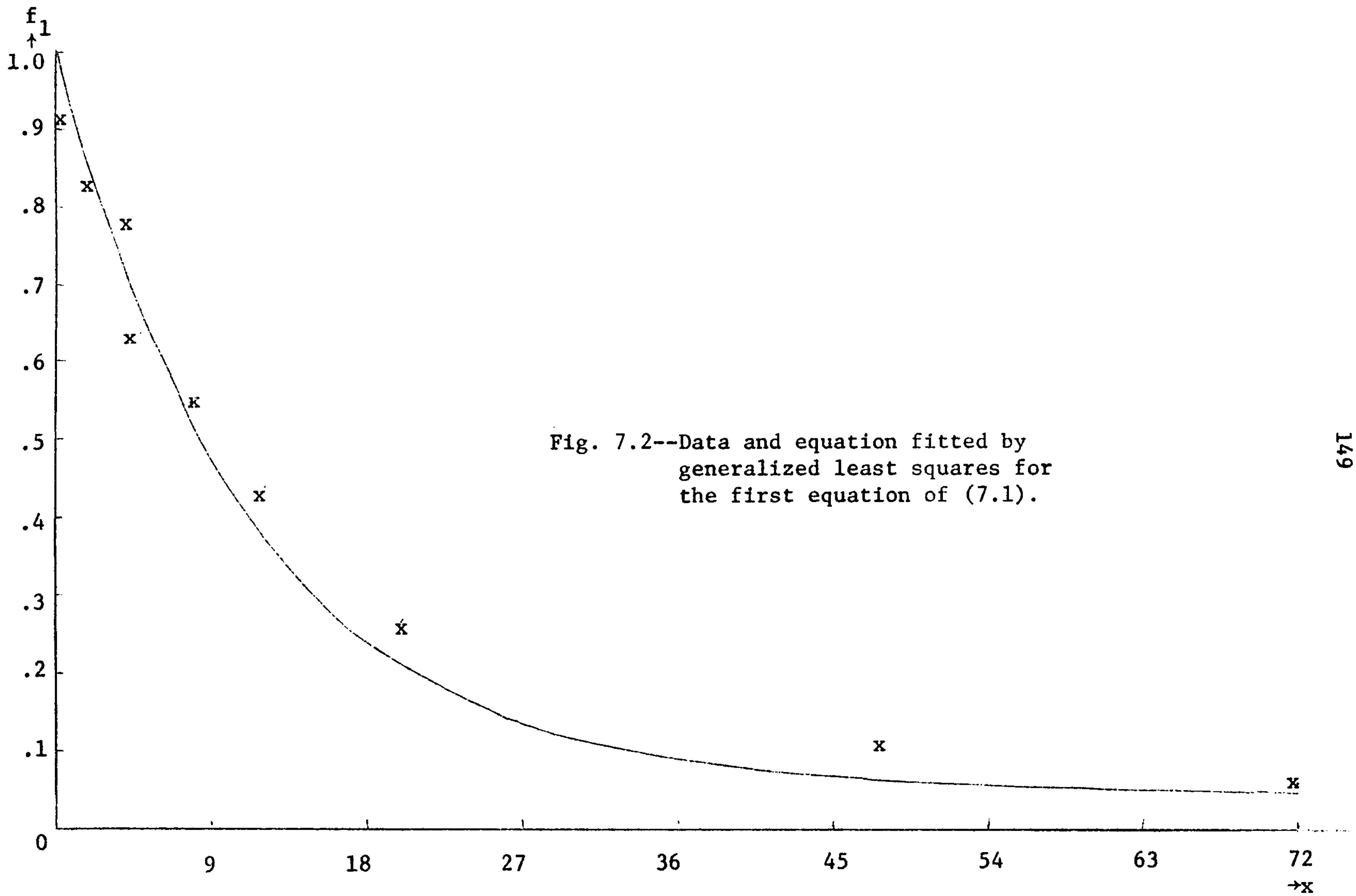

Fig. 7.2--Data and equation fitted by generalized least squares for the first equation of (7.1).

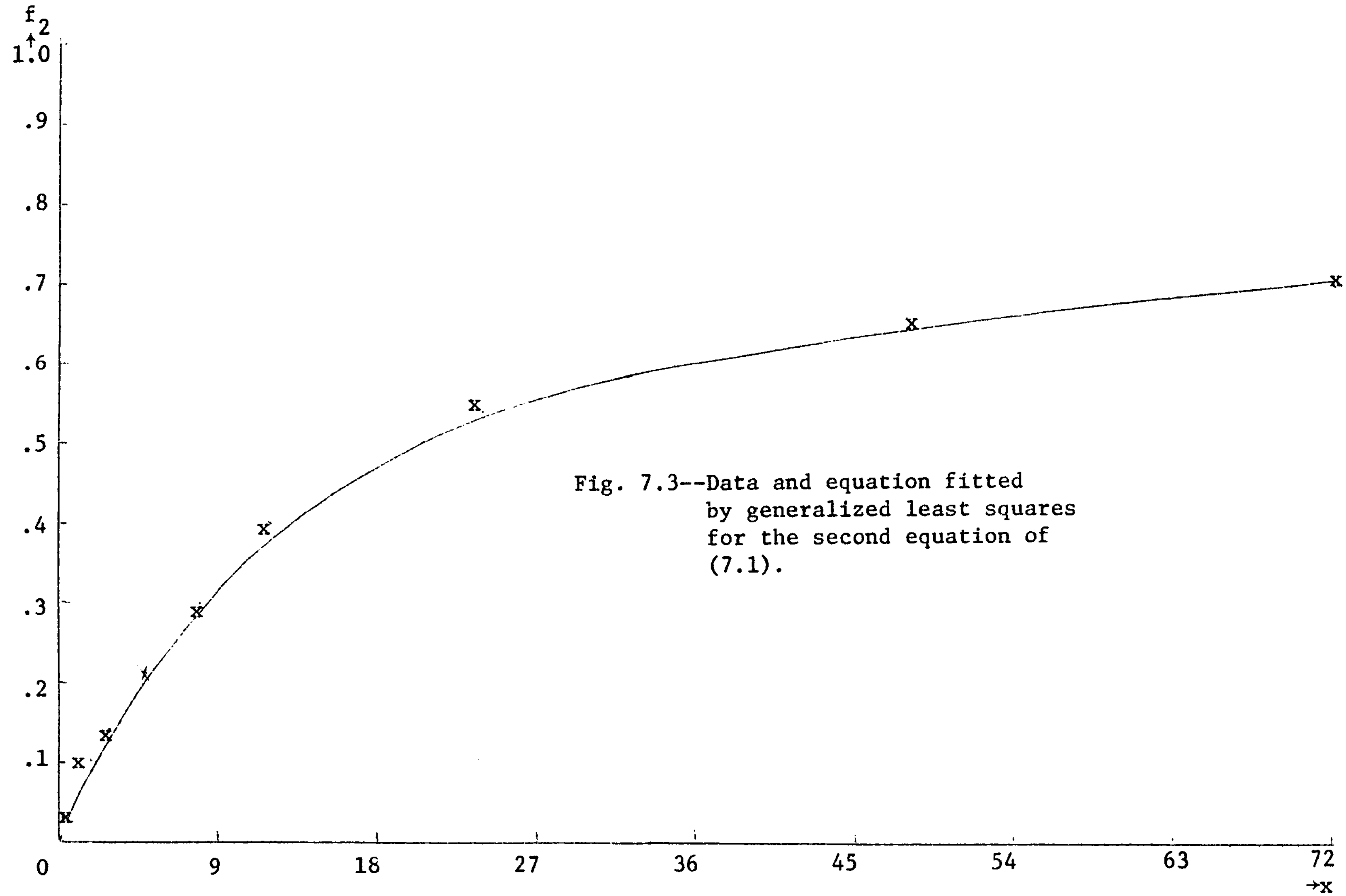

Fig. 7.3--Data and equation fitted by generalized least squares for the second equation of (7.1).

TABLE 7.3--Data to be fitted by generalized partial totals and least squares estimation procedures

$x_j(=j)$	y_{1j}	$y'_{2j}=1-y_{2j}$	Partial Totals
0	0.99580	0.98526	
1	0.86755	0.90118	S_{11} = 5.23783
2	0.75378	0.78387	
3	0.68462	0.72374	S_{21} = 5.63595
4	0.58998	0.64451	
5	0.49806	0.58602	
6	0.49066	0.57477	
7	0.35738	0.43660	
8	0.31896	0.44126	
9	0.32844	0.43487	S_{12} = 1.96023
10	0.24684	0.34459	
11	0.29593	0.38054	S_{22} = 2.80562
12	0.18045	0.28662	
13	0.25398	0.33810	
14	0.17297	0.29868	
15	0.16266	0.28096	
16	0.15076	0.24881	
17	0.12821	0.22204	S_{13} = 0.95647
18	0.12233	0.24219	
19	0.15341	0.29722	S_{23} = 1.82865
20	0.13334	0.24112	
21	0.08309	0.17590	
22	0.09083	0.19781	
23	0.09450	0.20356	

$$Y_{1j} = \alpha_1 e^{-\lambda_1 x_j} + (1-\alpha_1)e^{-\lambda_2 x_j} + \varepsilon_{1j}$$

$$Y_{2j} = 1 - \alpha_2 e^{-\lambda_1 x_j} - (1-\alpha_2)e^{-\lambda_2 x_j} + \varepsilon_{2j} \ . \tag{7.2}$$

Instead of recording the observations y_{2j} on the second regression equation of (7.2), we record $y'_{2j} = 1 - y_{2j}$ so that $E(Y'_{2j}) = \alpha_2 e^{-\lambda_1 x_j} + (1-\alpha_2)e^{-\lambda_2 x_j}$ will be of the same functional form as $E(Y_{ij})$.

For this particular example we have 24 observations on each regression equation, therefore we divide each set of observations up into three groups each containing eight observations and form the following partial totals:

$$S_{iq} = \sum_{j=(q-1)8}^{8q-1} y_{ij} \tag{7.3}$$

for $i = 1,2$ and $q = 1,2,3$. The values for these partial totals are also given in the last column of Table 7.3. Using equation (5.12) we are now able to use the set of equations given below to obtain estimates of the elementary symmetric functions of $e^{-8\lambda_1}$ and $e^{-8\lambda_2}$, denoted by L_1 and L_2:

$$S_{11}L_2 - S_{12}L_1 = -S_{13}$$

$$S_{21}L_2 - S_{22}L_1 = -S_{23} \tag{7.4a}$$

or

$$5.23783\ L_2 - 1.96023\ L_1 = -0.95647$$

$$5.63595\ L_2 - 2.80562\ L_1 = -1.82865\ . \tag{7.4b}$$

Solving the set of equations given in (7.4b), we find $L_1 = 1.14810$ and $L_2 = 0.24706$. In order to obtain the estimates of $e^{-8\lambda_1}$ and $e^{-8\lambda_2}$ we obtain the two roots of the following quadratic equation:

$$w^2 - 1.14810\ w + 0.24706 = 0\ . \tag{7.5}$$

The roots of (7.5) are given by $w_1 = 0.86123$ and $w_2 = 0.28687$. Without loss of generality we will assume that $\lambda_1 < \lambda_2$, and therefore our estimates of λ_1 and λ_2 are given by $\hat{\lambda}_1 = 0.01868 = \frac{-1}{8}\ln(0.86123)$ and $\hat{\lambda}_2 = \frac{-1}{8}\ln(0.28687) = 0.15609$, respectively.

The next step in our estimation procedure is to estimate the linear parameters, α_1 and α_2, present in our regression model given by equation (7.2). Our observations given in Table 7.3 were generated by adding random normal variables to calculated expected values and taking $E(\varepsilon_{1j}^2) = \sigma_{11} = 0.001$; $E(\varepsilon_{2j}^2) = \sigma_{22} = 0.001$; and $E(\varepsilon_{1j}\varepsilon_{2j}) = \sigma_{12} = 0.0009$ for all j.

Therefore in this particular example we may assume that the matrix Ω in Section 5.3 is known. Since there is only one linear parameter in each regression equation of our model, we may rewrite our regression model given in equation (7.2) as

$$Y_{1j} - e^{-\lambda_2 x_j} = \alpha_1(e^{-\lambda_1 x_j} - e^{-\lambda_2 x_j}) + \varepsilon_{1j}$$

$$1 - Y_{2j} - e^{-\lambda_2 x_j} = \alpha_2(e^{-\lambda_1 x_j} - e^{-\lambda_2 x_j}) + \varepsilon_{2j}. \qquad (7.6)$$

If λ_1 and λ_2 were known, then the usual weighted least squares estimators of α_1 and α_2 are given by

$$\hat{\alpha}_{**} = (D_{\tilde{Z}}^T \Omega^{-1} D_{\tilde{Z}})^{-1} (D_{\tilde{Z}}^T \Omega^{-1} \tilde{y}_{**}) \qquad (7.7)$$

where $\hat{\alpha}_{**} = (\hat{\alpha}_1, \hat{\alpha}_2)^T$, $D_{\tilde{Z}} = \begin{pmatrix} \tilde{Z} & 0 \\ 0 & \tilde{Z} \end{pmatrix}$;

$\tilde{Z} = (0, e^{-\lambda_1} - e^{-\lambda_2}, \ldots, e^{-23\lambda_1} - e^{-23\lambda_2})^T$; $\Omega = \sigma_{**} \otimes I$;

$\sigma_{**} = \begin{pmatrix} 0.001 & 0.0009 \\ 0.0009 & 0.001 \end{pmatrix}$; I is a 12×12 identity matrix; and

$\tilde{y}_{**} = (y_{10} - 1; y_{11} - e^{-\lambda_2}; \ldots; y_{1,23} - e^{-23\lambda_2}; -y_{20}; 1 - y_{21} - e^{-\lambda_2}; \ldots;$

$1 - y_{2,23} - e^{-23\lambda_2})^T$.

Using the same technique as we used to derive equation (5.17), we substitute the estimators $\hat{\lambda}_1$ and $\hat{\lambda}_2$ into equation (7.7) giving us the estimates of the parameters α_1 and α_2, denoted by $\hat{\hat{\alpha}}_1$ and $\hat{\hat{\alpha}}_2$, respectively. Applying the above procedure to this example we find $\hat{\hat{\alpha}}_1 = 0.12880$ and $\hat{\hat{\alpha}}_2 = 0.25528$.

In case we do not know the true values of σ_{11}, σ_{22}, and σ_{12}, we use the same method as we used to derive equation (5.17) to obtain estimates of σ_{11}, σ_{22}, and σ_{12}. Substituting $\hat{\lambda}_1$ and $\hat{\lambda}_2$ into the regression equations of our model, we first obtain the single equation least squares estimates of the parameters α_1 and α_2. These are given by $\hat{\alpha}_1 = 0.10900$ and $\hat{\alpha}_2 = 0.27569$. Using equation (5.17.1) we find $\hat{\sigma}_{11} = 0.00064$; $\hat{\sigma}_{22} = 0.00074$; and $\hat{\sigma}_{12} = 0.00058$. Substituting these estimates into the matrix Ω in equation (7.7) we find our new estimates of α_1 and α_2 given by 0.14735 and 0.27007, respectively.

For comparison, we will obtain the generalized least squares estimates of the parameters λ_1, λ_2, α_1, and α_2. Since we have already outlined the basic steps of the generalized least squares procedure in Section 7.2.1, we merely give the results here. The complete results of the various calculations in this section are summarized in Table 7.4. Across the top of Table 7.4 are listed the various estimation techniques used to obtain the values listed in the table and down the first four positions of the first column of this table are listed the various parameters of interest. In the last two rows of this table are listed the

TABLE 7.4--Estimates of parameters in equation (7.2)

Parameter	Generalized Partial Totals		Generalized Least Squares	
	Ω Known	Ω Unknown	Ω Known	Ω Unknown
α_1	0.12880	0.14735	0.09461	0.09504
α_2	0.25528	0.27007	0.25409	0.25388
λ_1	0.01868	0.01868	0.01546	0.01528
λ_2	0.15609	0.15609	0.15229	0.15261
Figure Number	7.4	7.6	7.8	7.10
Figure Number	7.5	7.7	7.9	7.11

various figures where graphs of the original data and the fitted regression equations of our model appear for a visual comparison.

7.2.3 Generalized Spearman estimation

In this section we apply the generalized Spearman estimation technique, developed in Chapter 6, to the set of data given in Table 7.5 generated from an experimental situation similar to that considered by Box and Draper [1965]. The experimental situation of interest involves a chemical reaction in which a product 3 is decomposing to form product 2 which in turn decomposes to form product 1. Schematically we may represent this chemical reaction

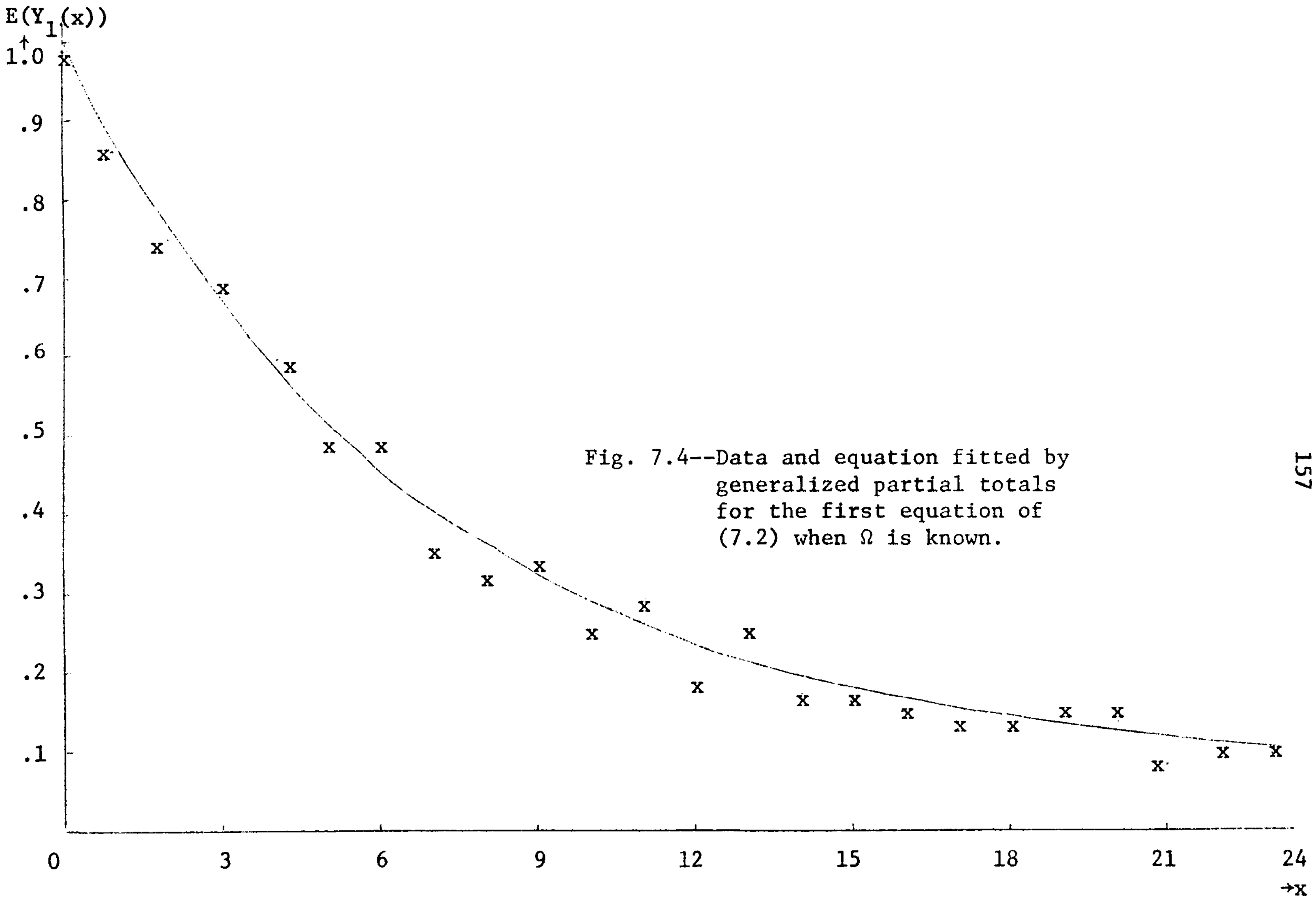

Fig. 7.4--Data and equation fitted by generalized partial totals for the first equation of (7.2) when Ω is known.

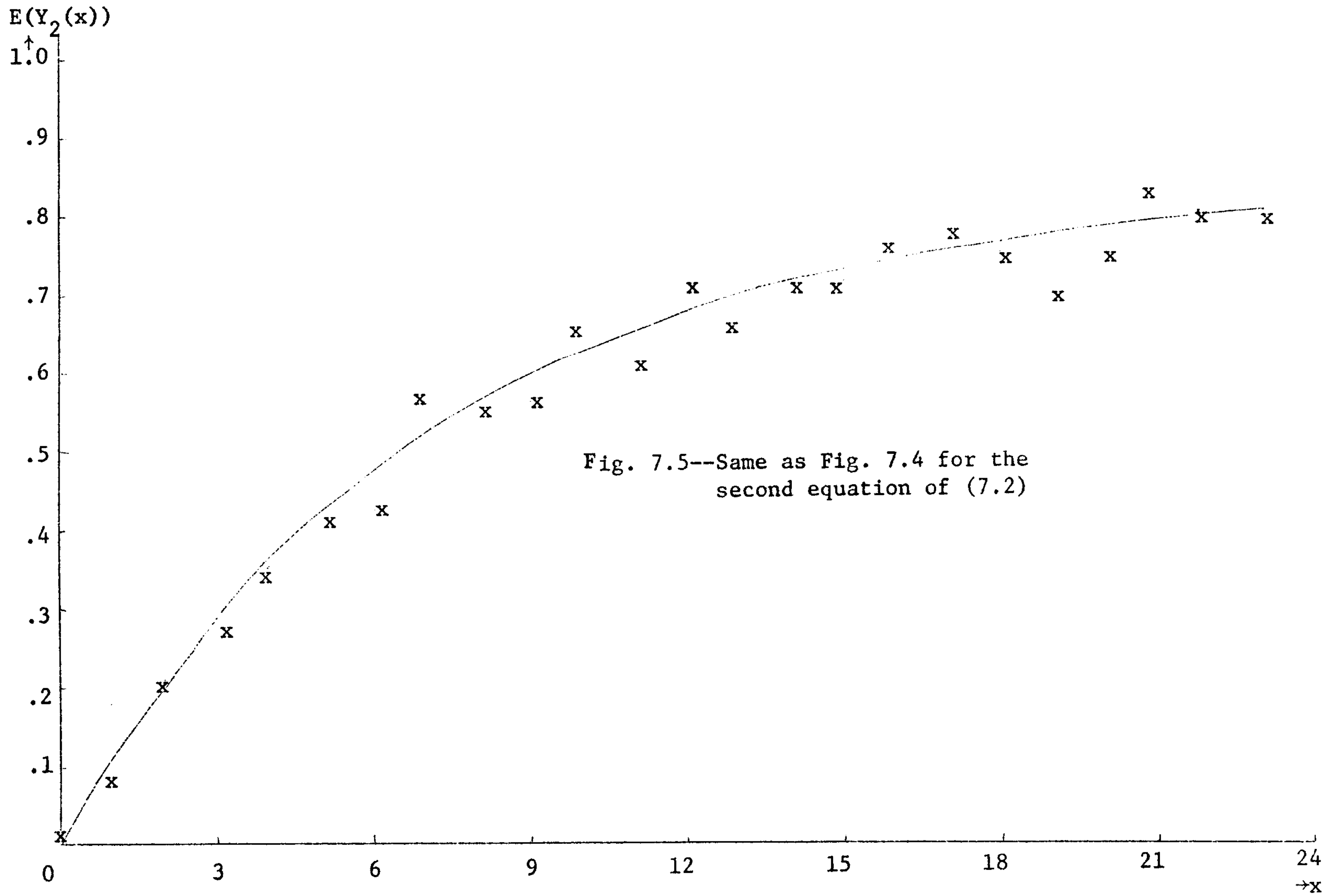

Fig. 7.5--Same as Fig. 7.4 for the second equation of (7.2)

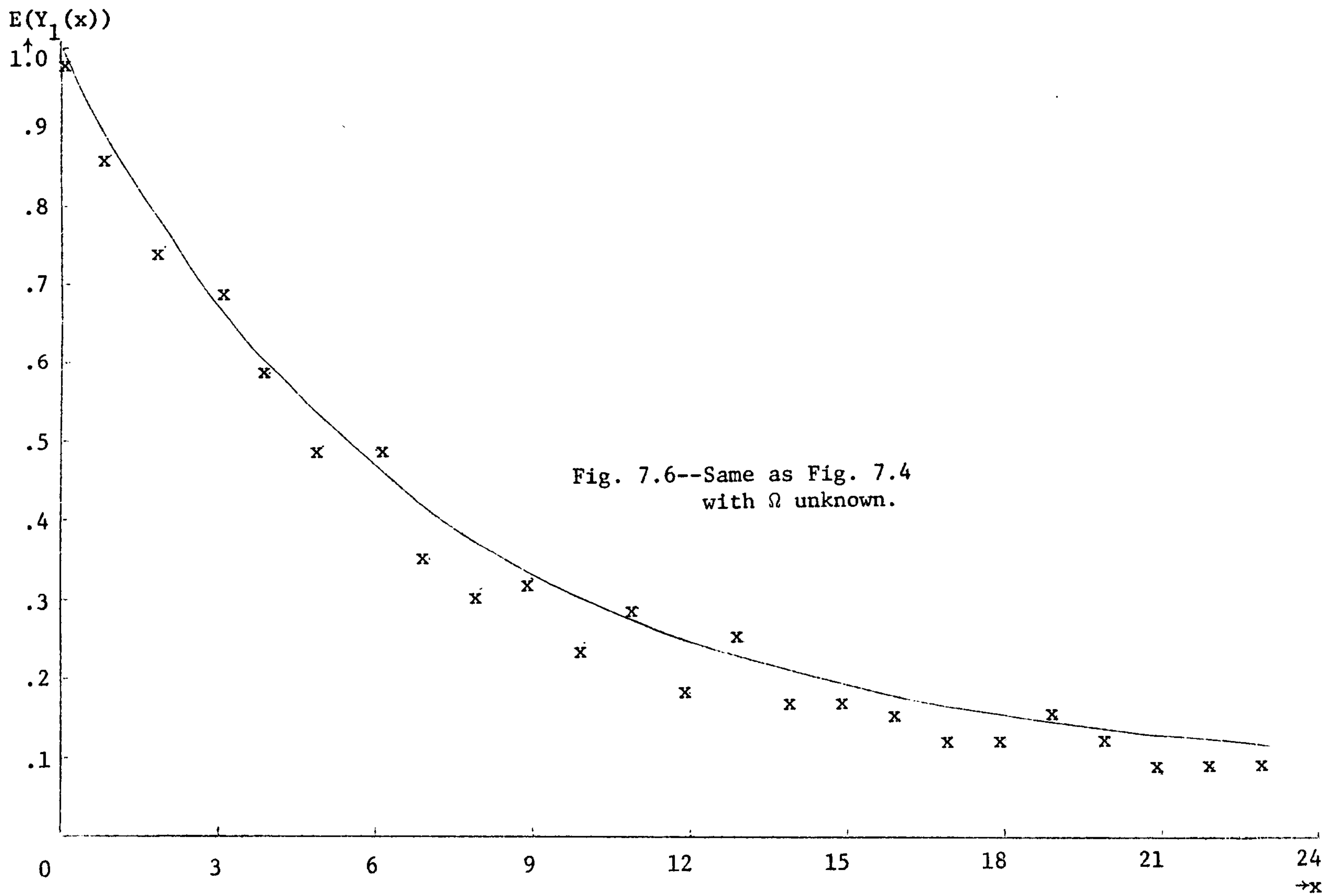

Fig. 7.6--Same as Fig. 7.4 with Ω unknown.

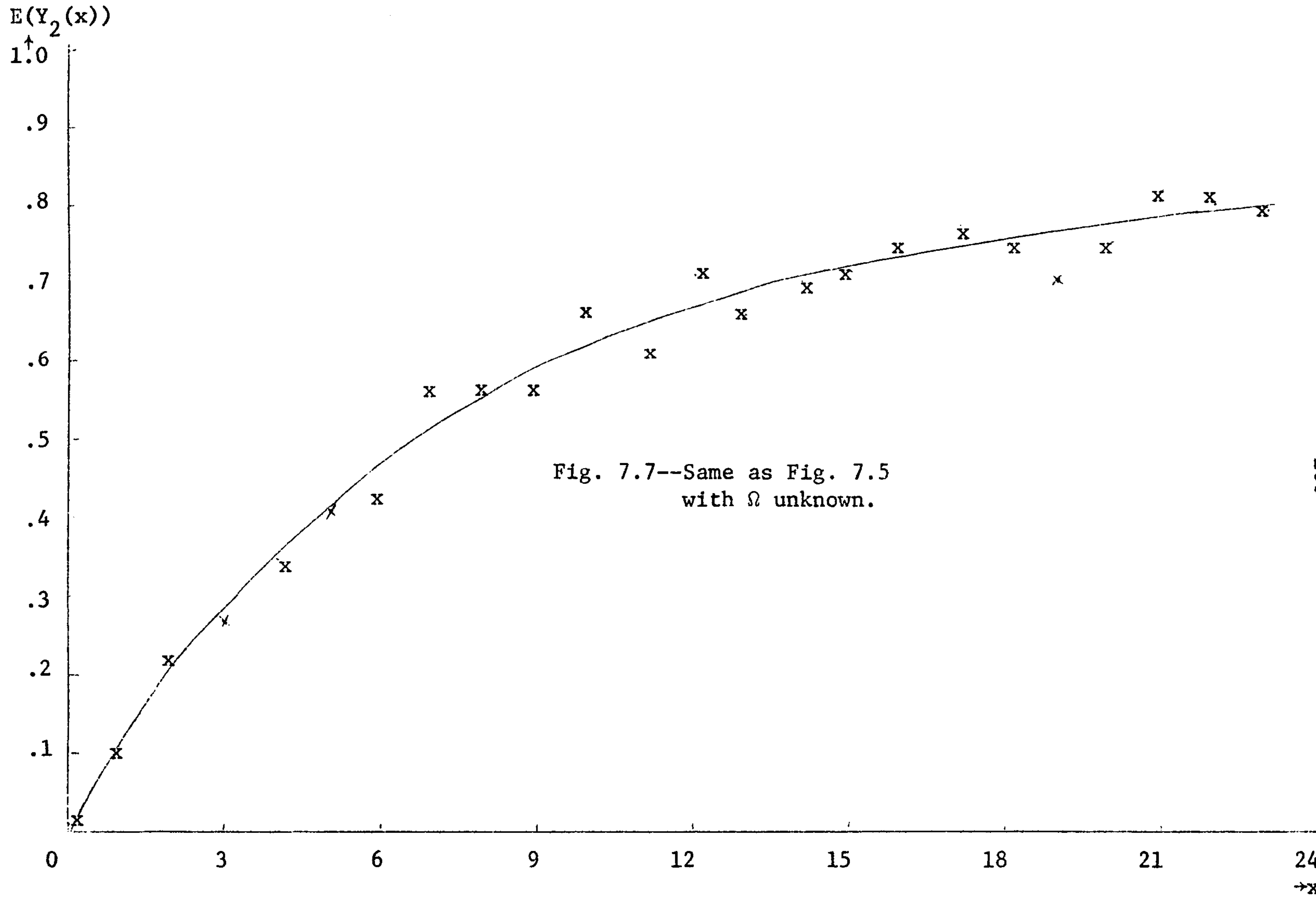

Fig. 7.7--Same as Fig. 7.5 with Ω unknown.

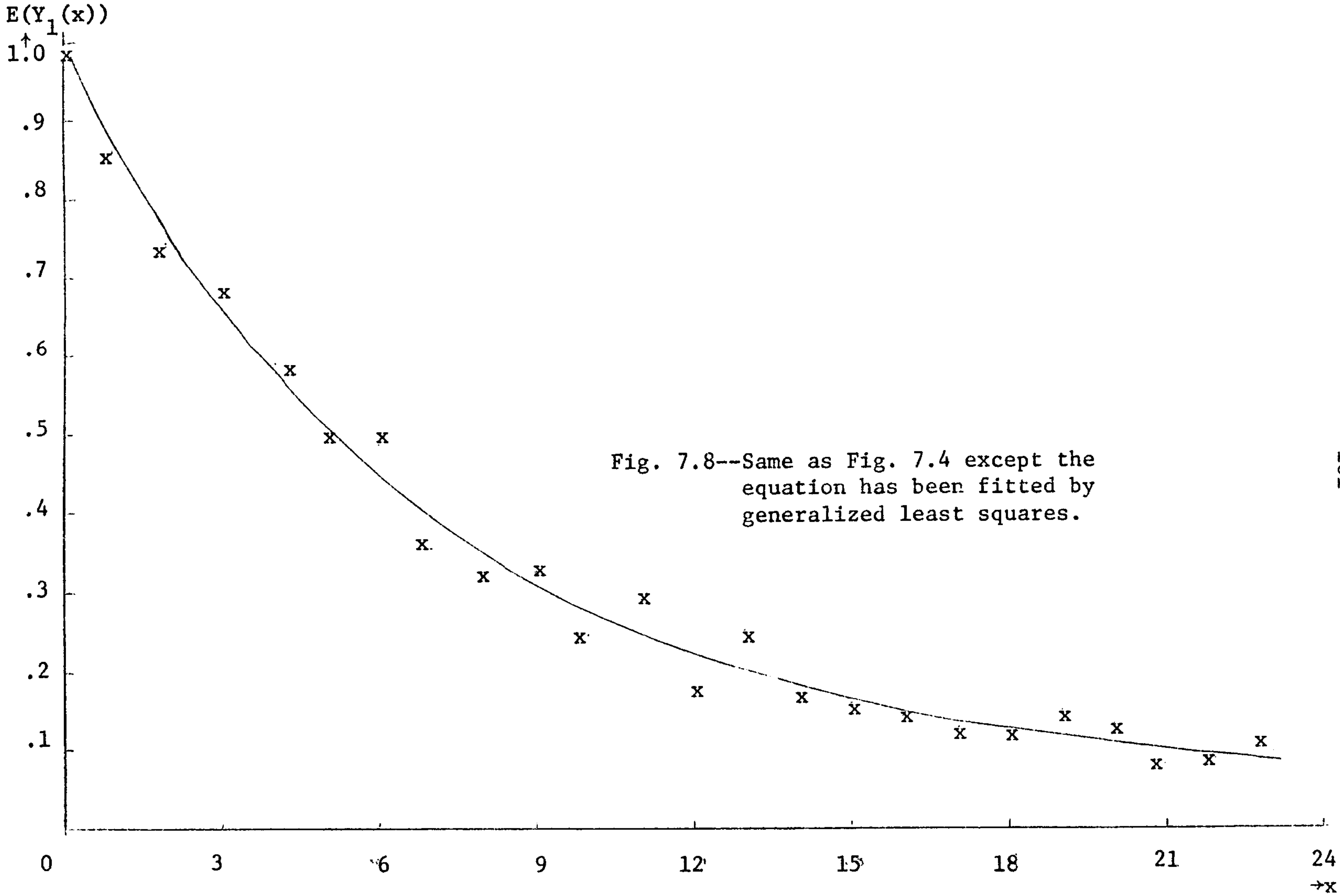

Fig. 7.8--Same as Fig. 7.4 except the equation has been fitted by generalized least squares.

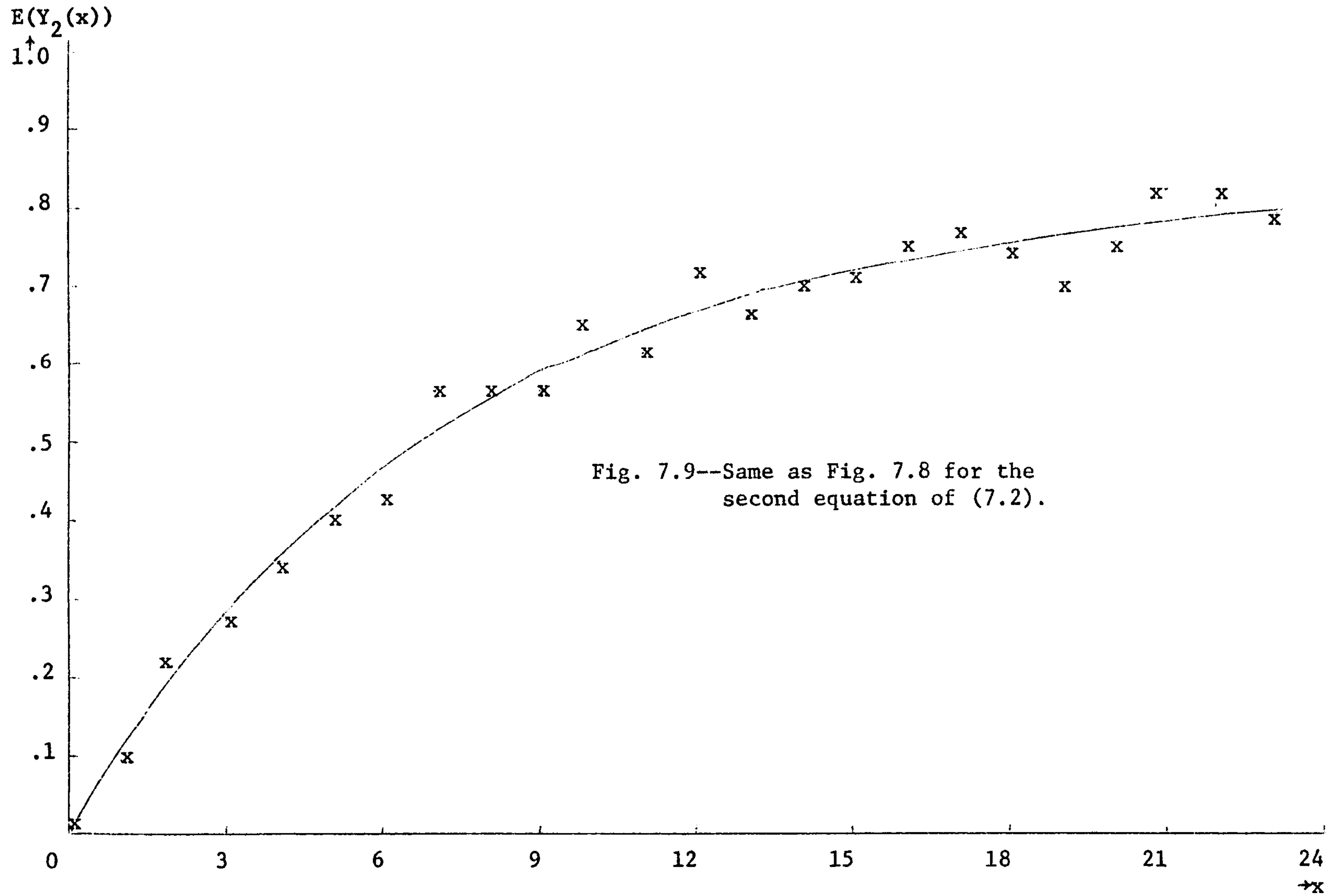

Fig. 7.9--Same as Fig. 7.8 for the second equation of (7.2).

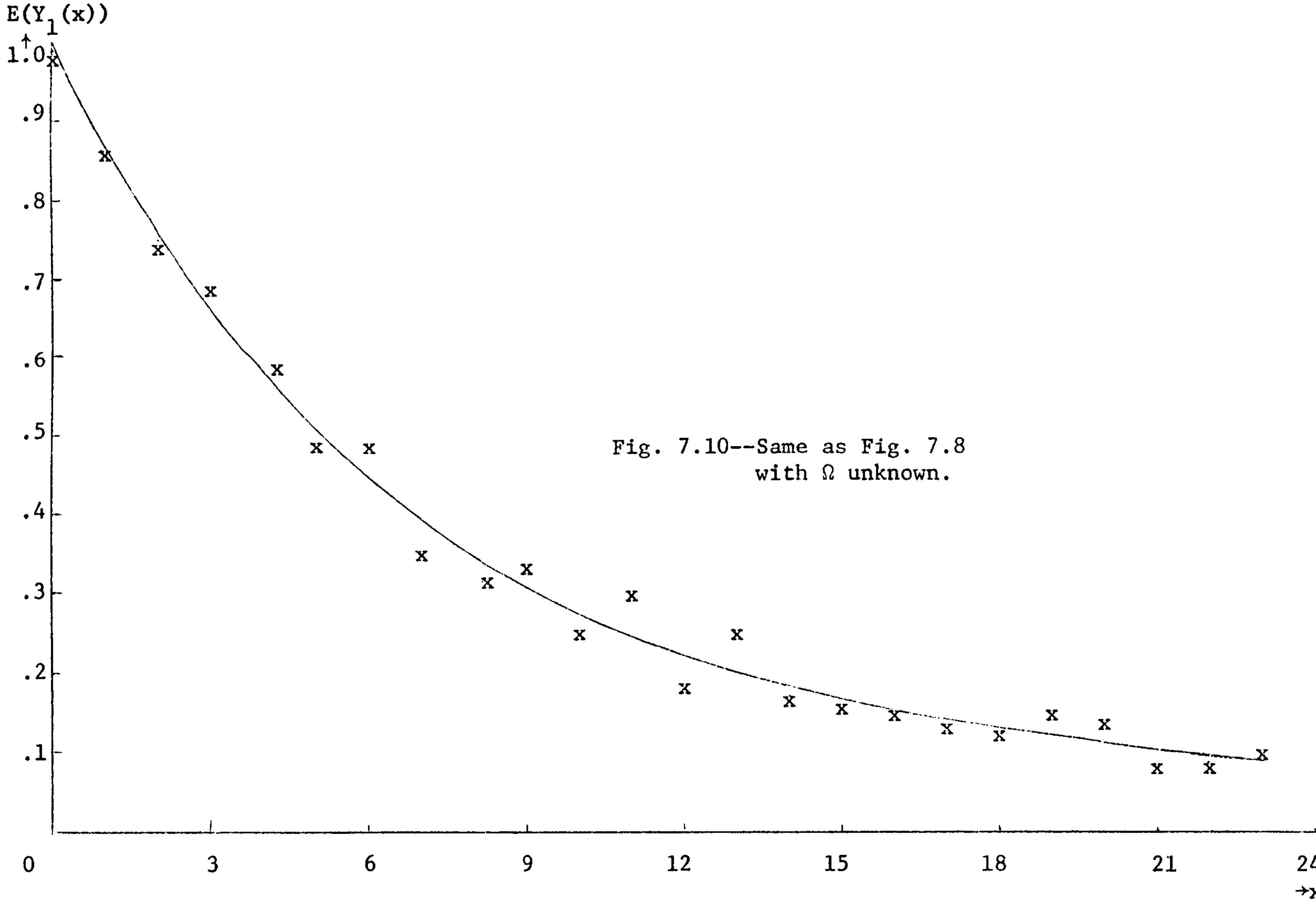

Fig. 7.10--Same as Fig. 7.8 with Ω unknown.

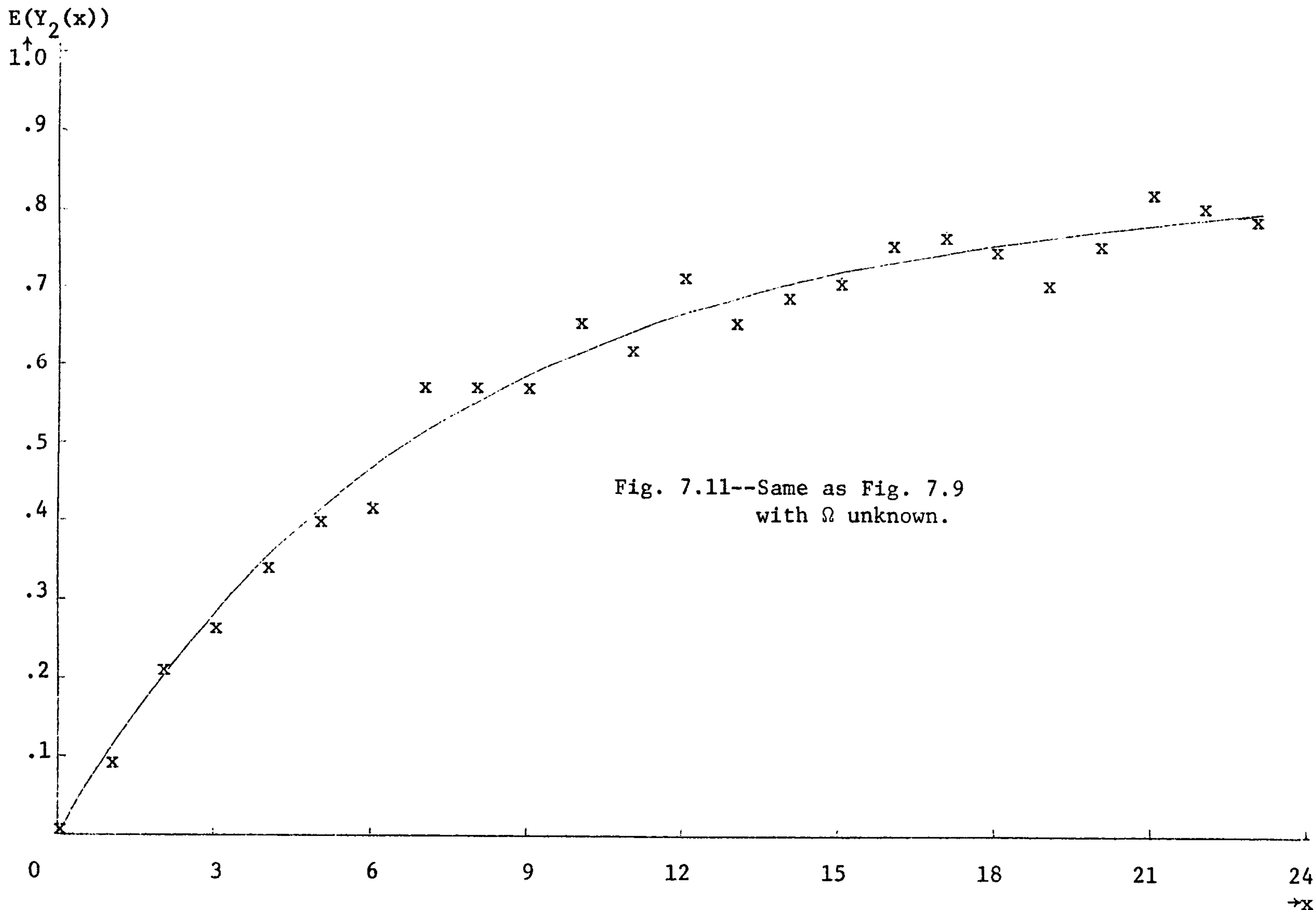

Fig. 7.11--Same as Fig. 7.9
with Ω unknown.

by the following diagram:

$$\boxed{3} \xrightarrow{\tau_{32}} \boxed{2} \xrightarrow{\tau_{12}} \boxed{1}$$

TABLE 7.5--Data to be fitted by generalized Spearman and least squares estimation procedures

j	x_j	z_j	y'_{1j}	y_{2j}
-5	$\frac{1}{4}$	-1.38630	1.00000	0.00000
-4	$\frac{1}{2}$	-0.69315	0.92696	0.01463
-3	1	0	0.87213	0.02986
-2	2	0.69315	0.75029	0.04675
-1	4	1.38630	0.60339	0.10608
0	8	2.07945	0.37711	0.16495
1	16	2.77260	0.18042	0.19098
2	32	3.46575	0.05943	0.14191
3	64	4.15890	0.02628	0.09154
4	128	4.85205	0.00000	0.00000

In the above diagram τ_{23} and τ_{12} represent the reaction rates from product 3 to product 2 and from product 2 to product 1, respectively. We now note that we have brought this problem into the same framework as the compartmental problem discussed in Chapter 3. Therefore we have a system of differential equations corresponding to those given in equation (3.1) for n = 2. For this particular example let $E(Y_i(x))$, i = 1,2,3, represent the expected proportion of product i present at time x and let the following boundary conditions be satisfied:

1) $E(Y_3(x)) = 1$ at $x = 0$; and 2) $E(Y_2(x)) = E(Y_1(x)) = 0$ at $x = 0$. Corresponding to equation (3.7) we have the following system of equations:

$$E(Y_1(x)) = \alpha_{11}e^{-\lambda_1 x} + \alpha_{12}e^{-\lambda_2 x} + \alpha_{13}e^{-\lambda_3 x}$$

$$E(Y_2(x)) = \alpha_{21}e^{-\lambda_1 x} + \alpha_{22}e^{-\lambda_2 x} + \alpha_{23}e^{-\lambda_3 x}$$

$$E(Y_3(x)) = \alpha_{31}e^{-\lambda_1 x} + \alpha_{32}e^{-\lambda_2 x} + \alpha_{33}e^{-\lambda_3 x} \tag{7.8}$$

where $\lambda_1 = \tau_{23}$; $\lambda_2 = \tau_{12}$; $\lambda_3 = 0$; $\alpha_{11} = \dfrac{-\tau_{12}}{\tau_{12}-\tau_{23}}$; $\alpha_{12} = 1 - \alpha_{11}$;

$\alpha_{13} = 1$; $\alpha_{21} = -\alpha_{22} = \dfrac{\tau_{23}}{\tau_{12}-\tau_{23}}$; $\alpha_{23} = \alpha_{32} = \alpha_{33} = 0$; and

$\alpha_{31} = -\tau_{23}$. Since $E(Y_1(x))$, $E(Y_2(x))$, and $E(Y_3(x))$ represent the expected proportions of the various products present at time x, we have the additional restriction that $\sum_{i=1}^{3} E(Y_i(x)) = 1$ for all x. Hence there are only two independent equations in (7.8), which we take to be the first two. If we let $\alpha_1 = -\alpha_{11}$, $x = e^z$, and

$E(Y_1'(x)) = 1 - E(Y_1(x))$, then the equation for $E(Y_1'(x))$ becomes

$$E(Y_1'(z)) = \alpha_1 \exp(-\lambda_1 e^z) + (1-\alpha_1)\exp(-\lambda_2 e^z) \tag{7.9a}$$

which is of the same form as equation (6.6a). In addition, if we let $\alpha_2 = \alpha_{21}$ then the equation for $E(Y_2(x))$ is given by

$$E(Y_2(z)) = \alpha_2 \exp(-\lambda_1 e^z) - \alpha_2 \exp(-\lambda_2 e^z) \tag{7.9b}$$

which is of the same form as equation (6.6b).

By the evaluation of the integrals $\mu_1^{(k')} = \int_{-\infty}^{\infty} z^{k'} dE(Y_1'(z))$ and $\mu_2^{(k')} = \int_{-\infty}^{\infty} z^{k'} dE(Y_2(z))$ for $k' = 1,2$, and by using the same techniques as presented in Section 6.3, we arrive at the system of equations corresponding to (6.25b) and (6.33e) given by

$$\begin{aligned} K_{11}\Lambda_1^{(C)} - \Lambda_2^{(C)} &= - K_{12} \\ K_{21}'\Lambda_1^{(C)} \qquad &= - K_{22}' \end{aligned} \tag{7.10a}$$

where $\Lambda_1^{(C)} = \ln\lambda_1 + \ln\lambda_2$; $\Lambda_2^{(C)} = \ln\lambda_1 \ln\lambda_2$; $K_{11} = \mu_1^{(1)} + I_1$; $K_{12} = \mu_1^{(2)} + I_2 - 2I_1K_{11}$; $K_{21}' = \mu_2^{(1)}$; $K_{22}' = \mu_2^{(2)} - 2I_1K_{21}'$;

$I_1 = \int_0^{\infty} (\ln t)e^{-t}dt$; and $I_2 = \int_0^{\infty} (\ln t)^2 e^{-t}dt$.

All that we now need to obtain our estimates of the elementary symmetric functions $\Lambda_1^{(C)}$ and $\Lambda_2^{(C)}$, are the estimates of

$\mu_1^{(1)}$, $\mu_1^{(2)}$, $\mu_2^{(1)}$, and $\mu_2^{(2)}$. These estimates are found from equation (6.35a), and the values for this particular example are given by $\hat{\mu}_1^{(1)} = -1.59148$; $\hat{\mu}_1^{(2)} = -4.28851$; $\hat{\mu}_2^{(1)} = -0.54531$; and $\hat{\mu}_2^{(2)} = -2.65392$. Using these calculated values to obtain our estimates of the K's, the system of equations given in (7.10a) becomes

$$\begin{aligned} -2.16870\, L_1^{(C)} - L_2^{(C)} &= 4.81403 \\ -0.54531\, L_1^{(C)} &= 3.28345 \ . \end{aligned} \tag{7.10b}$$

Solving the set of equations given in (7.10b), we find $L_1^{(C)} = -6.02125$ and $L_2^{(C)} = 8.24425$. In order to obtain the estimates of $\ln\lambda_1$ and $\ln\lambda_2$ we obtain the two roots of the following quadratic equation:

$$w^2 + 6.02125\, w + 8.24425 = 0 \ . \tag{7.11}$$

The roots of (7.11) are given by $w_1 = -3.91595$ and $w_2 = -2.10530$. Without loss of generality we assume that $\lambda_1 < \lambda_2$, and therefore our estimates of λ_1 and λ_2 are given by $\hat{\lambda}_1 = e^{-3.91595} = 0.01990$ and $\hat{\lambda}_2 = e^{-2.10530} = 0.12181$, respectively.

The next step in our estimation procedure will be to estimate the linear parameters, α_1 and α_2, present in our regression model given by equation (7.2). Since we generated the observations given in Table 7.5, we know the covariance matrix of the random variables ε_{1j} and ε_{2j}. Using the observations from Table 7.5 for $j = -4$ through $j = 3$, we use a weighted least squares procedure to estimate the parameters α_1 and α_2. We merely give the results of this estimation here, since the procedure used to estimate α_1 and α_2 has already been outlined in Section 7.2.2. In order to obtain a comparison of the estimation procedures, we also obtain the generalized least squares estimates of the parameters λ_1, λ_2, α_1, and α_2 for this example. The complete results of the various calculations in this section are summarized in Table 7.6, which is arranged in the same manner as Table 7.4.

TABLE 7.6--Estimates of parameters in equation (7.8)

Parameter	Generalized Spearman	Generalized Least Squares
α_1	0.05501	0.09231
α_2	0.30899	0.27832
λ_1	0.01990	0.01813
λ_2	0.12181	0.14239
Figure Number	7.12	7.14
Figure Number	7.13	7.15

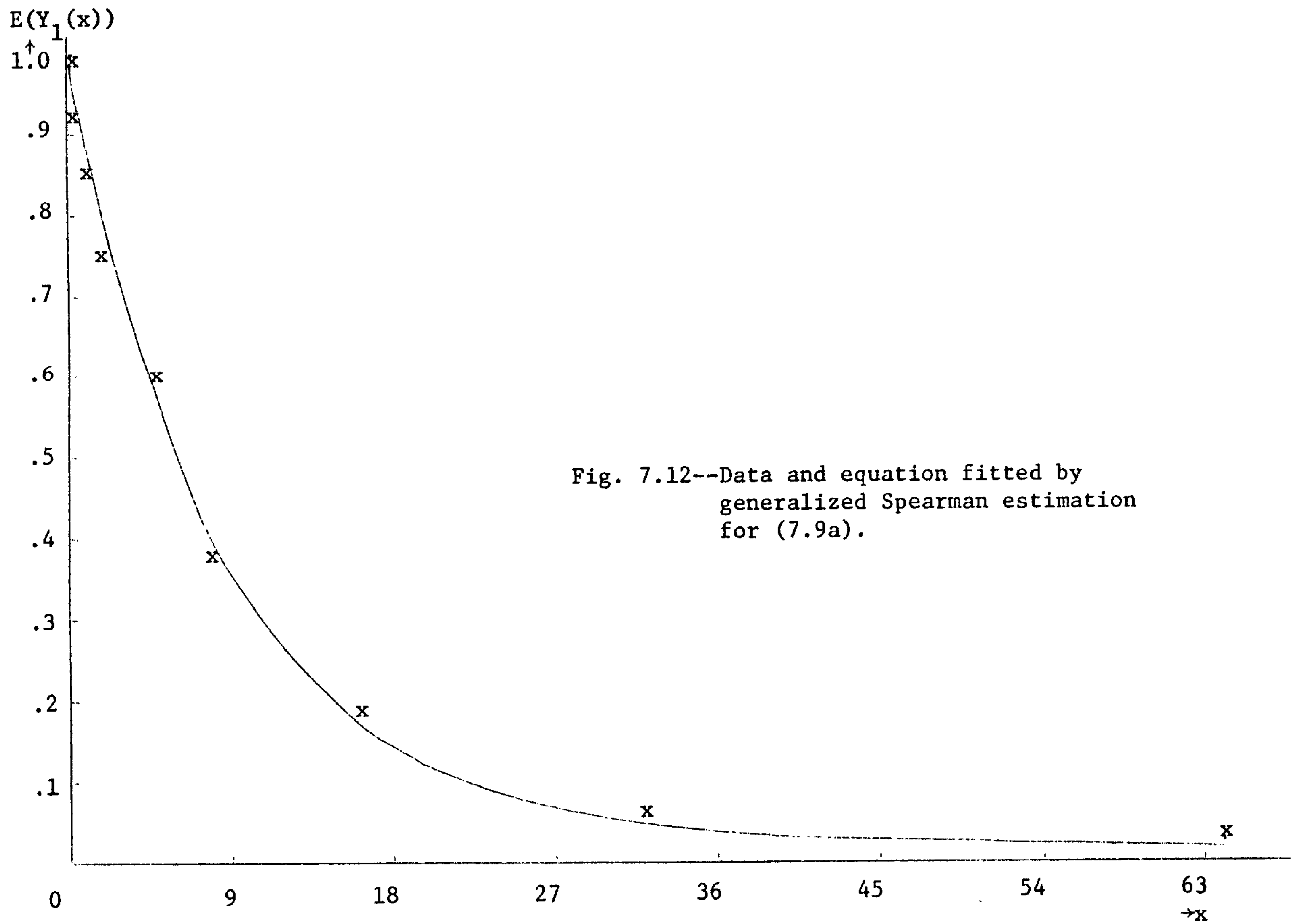

Fig. 7.12--Data and equation fitted by generalized Spearman estimation for (7.9a).

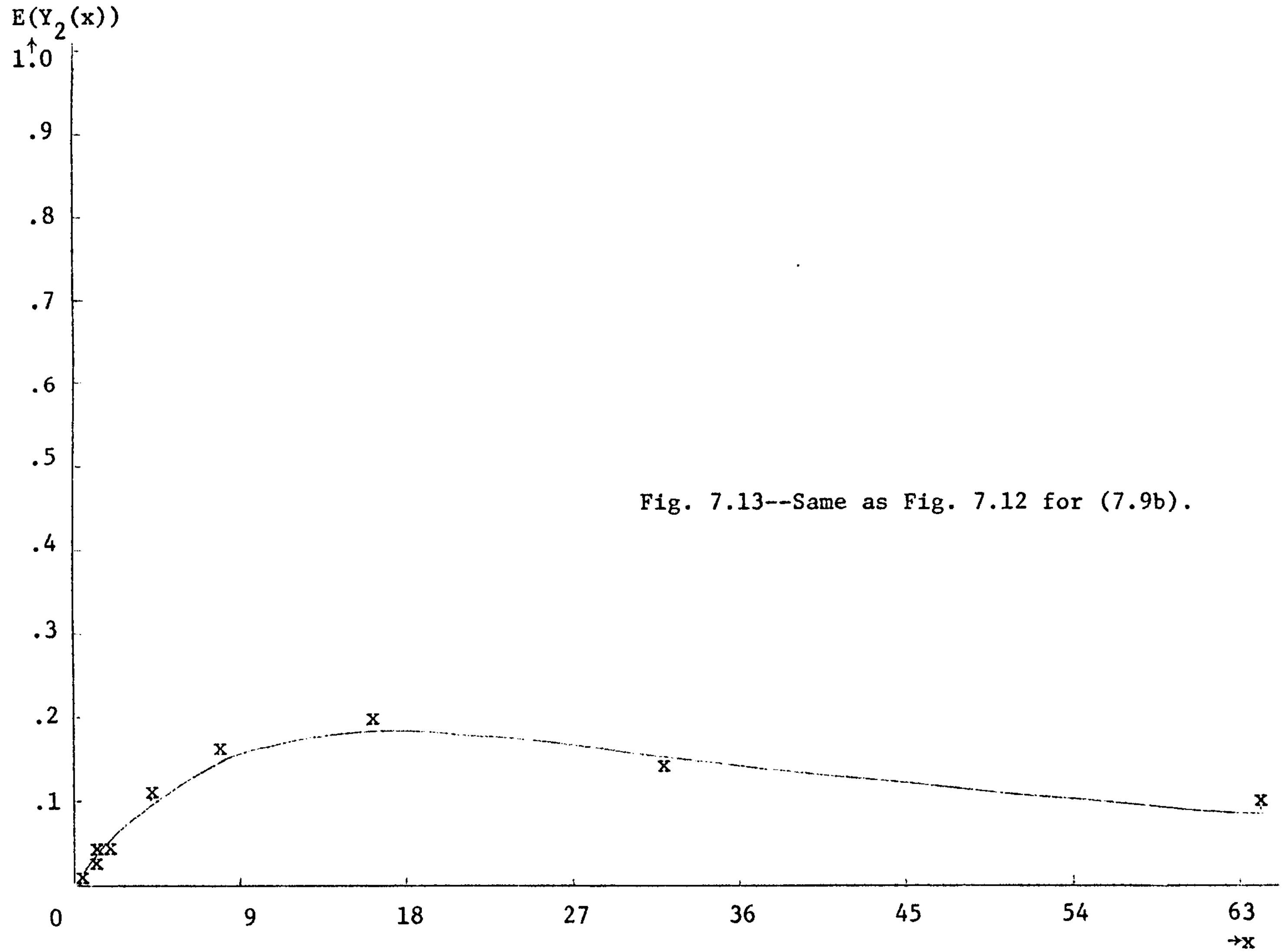

Fig. 7.13--Same as Fig. 7.12 for (7.9b).

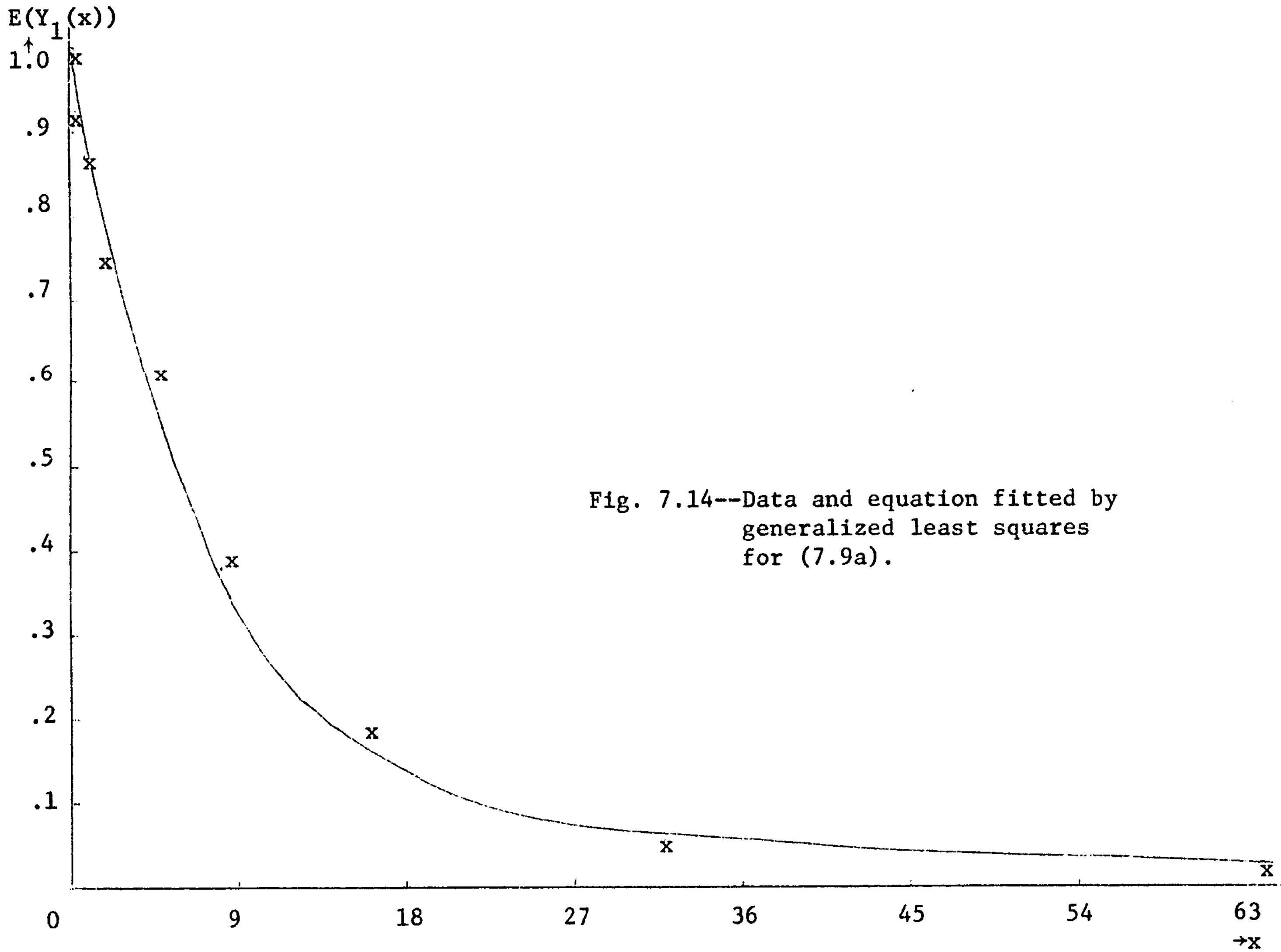

Fig. 7.14--Data and equation fitted by generalized least squares for (7.9a).

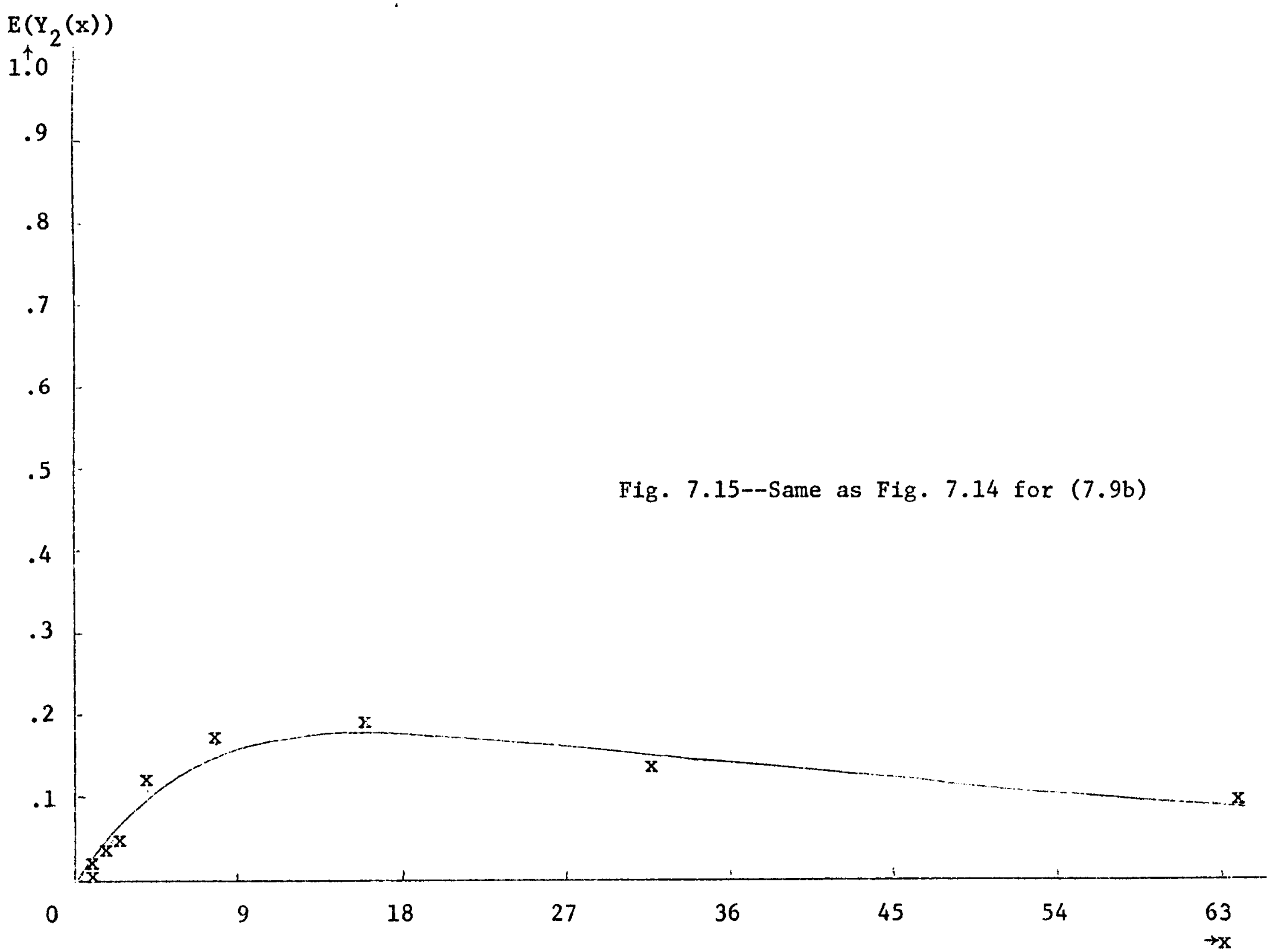

Fig. 7.15--Same as Fig. 7.14 for (7.9b)

7.3 Evaluation of the asymptotic efficiency for various models

7.3.1 Generalized partial totals estimation procedure

In Section 5.4 we stated that we would evaluate the asymptotic efficiency of the generalized partial totals estimators of the exponential parameters for some particular regression models. The expression for the asymptotic efficiency of the exponential parameters has been defined earlier in equation (5.42). In this section we evaluate this expression for some particular regression models of interest.

First let us consider the regression model given by

$$Y_{1j} = \alpha_1 e^{-\lambda_1 x_j} + \varepsilon_{1j}$$

for $j = 0,1,2,\ldots,2M-1$. In addition, we assume that the random variables ε_{1j} are independent each with a $N(0,\sigma^2)$ distribution. With these assumptions we have satisfied the assumptions about the random variables given at the beginning of Section 5.4, where some of the asymptotic properties of our generalized partial totals estimators were discussed. The likelihood function for our observed random variables Y_{1j} is given by

$$L = \prod_{j=o}^{2M-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{1}{2\sigma^2}(Y_{1j}-\alpha_1 e^{-\lambda_1 x_j})^2 \right) . \qquad (7.11)$$

We observe that

$$\frac{\partial \ln L}{\partial \lambda_1} = -\frac{\alpha_1}{\sigma^2} \sum_{j=o}^{2M-1} x_j e^{-\lambda_1 x_j} (Y_{1j} - \alpha_1 e^{-\lambda_1 x_j}), \tag{7.12}$$

and we can demonstrate that

$$E\left(\frac{\partial \ln L}{\partial \lambda_1}\right)^2 = \frac{\alpha_1^2}{\sigma^2} \sum_{j=o}^{2M-1} x_j^2 e^{-2\lambda_1 x_j} \; . \tag{7.13}$$

Using the results from Section 5.3, we recall that $x_j = hj$ and find the generalized partial totals estimator for λ_1, denoted by $\hat{\lambda}_1$, given by

$$\hat{\lambda}_1 = \frac{1}{J} \ln\left(\overline{S}_{11} / \overline{S}_{12}\right) \tag{7.14}$$

where $\overline{S}_{11} = \frac{1}{M} \sum_{j=o}^{M-1} y_{1j}$, $\overline{S}_{12} = \frac{1}{M} \sum_{j=M}^{2M-1} y_{1j}$, and $J = hM$ is the constant length of the domain of each partial total mentioned in the third assumption of Theorem 5.1. Using the results from Theorem 5.5, we find the asymptotic variance of the generalized partial totals estimator of λ_1 given by

$$\mathrm{Var}(\hat{\lambda}_1) = \frac{1}{M} F\Omega F^T \; . \tag{7.15}$$

For this particular example $\Omega = ME(\overline{\varepsilon}_{*}.\overline{\varepsilon}_{*}^{T}.)$; $\overline{\varepsilon}_{*}. = \frac{1}{M} \sum_{j=o}^{M-1} \varepsilon_{*j}$;

$$\varepsilon_{*j} = (\varepsilon_{1j}, \varepsilon_{1,j+M})^T;\text{ and } F = \left(\frac{\partial \hat{\lambda}_1}{\partial \bar{S}_{11}} \Big|_{\bar{S}_{**}=\Psi_{**}} , \frac{\partial \hat{\lambda}_1}{\bar{S}_{12}} \Big|_{\bar{S}_{**}=\Psi_{**}} \right).$$

Hence $\Omega = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$, and since $\Psi_{**} = \left(\frac{\alpha_1(1-e^{-\lambda_1 J})}{\lambda_1 J} ; \frac{\alpha_1 e^{-\lambda_1 J}(1-e^{-\lambda_1 J})}{\lambda_1 J} \right)^T$

we note that $F = \left(\frac{\lambda_1}{\alpha_1(1-e^{-\lambda_1 J})} , \frac{-\lambda_1}{\alpha_1 e^{-\lambda_1 J}(1-e^{-\lambda_1 J})} \right)$. Combining these results we find

$$\frac{1}{M} F\Omega F^T = \frac{\sigma^2 \lambda_1^2 (1+e^{2\lambda_1 J})}{M\alpha_1^2 (1-e^{-\lambda_1 J})^2} . \tag{7.16}$$

Next consider the product

$$E\left(\frac{\partial \ln L}{\partial \lambda_1}\right)^2 \text{Var}(\hat{\lambda}_1) = \frac{\lambda_1^2 (1+e^{2\lambda_1 J})}{(1-e^{-\lambda_1 J})^2} \left\{ \frac{1}{M} \sum_{j=o}^{2M-1} (hj)^2 \, e^{-2\lambda_1 h_j} \right\} . \tag{7.17}$$

The last term in braces is the only one that depends upon M, and from the assumptions of Theorem 5.1 and the definition of a definite integral we have

$$\lim_{M\to\infty} \frac{1}{M} \sum_{j=o}^{2M-1} (hj)^2 e^{-2\lambda_1 hj} = \frac{1}{J} \lim_{M\to\infty} \sum_{j=o}^{2M-1} \left(\frac{J}{M} j\right)^2 e^{-2\lambda_1 \frac{J}{M} j} \frac{J}{M}$$

$$= \frac{1}{J} \int_o^{2J} t^2 e^{-2\lambda_1 t} dt = \frac{1}{4J\lambda_1^3} (1-e^{-4\lambda_1 J} - 4\lambda_1 J e^{-4\lambda_1 J} - 8J^2\lambda_1^2 e^{-4\lambda_1 J}) . \tag{7.18}$$

Using the above results and the definition of the efficiency, ν, of the estimator $\hat{\lambda}_1$ given in equation (5.42), we have the following expression for this efficiency:

$$\nu = \frac{4\lambda_1 J(1-e^{-\lambda_1 J})^2}{(1+e^{2\lambda_1 J})(1-e^{-4\lambda_1 J} - 4\lambda_1 J e^{-4\lambda_1 J} - 8\lambda_1^2 J^2 e^{-4\lambda_1 J})} . \tag{7.19}$$

It is interesting to note that the expression for ν is independent of the parameter α_1 and is a function of $\lambda_1 J$ only. As $\lambda_1 J \to \infty$ it is easily seen that $\nu \to 0$; and as $\lambda_1 J \to 0$, by a repeated application of L'Hôpital's rule, it can be shown that $\nu \to 0.1875$. In Table 7.7 we have given the value of ν for various values of $\lambda_1 J$. By an examination of Table 7.7 we can see, for this particular model, that the efficiency of our generalized partial totals estimator of the parameter λ_1 achieves its maximum value around $\lambda_1 J = 0.7$. In addition, the efficiency does not vary a great deal for values of $\lambda_1 J$ between 0.5 and 1.0.

TABLE 7.7--Values of asymptotic efficiency of the partial totals estimator of the exponential parameter

$\lambda_1 J$	ν	$\lambda_1 J$	ν
0.0001	0.18753	0.70	0.26457
0.001	0.18769	0.71	0.26456
0.05	0.19676	0.75	0.26414
0.10	0.20573	0.80	0.26287
0.15	0.21432	0.85	0.26078
0.20	0.22245	0.90	0.25792
0.25	0.23004	0.95	0.25433
0.30	0.23702	1.00	0.25006
0.35	0.24331	1.20	0.22736
0.40	0.24886	1.40	0.19857
0.45	0.25362	1.60	0.16742
0.50	0.25756	1.80	0.13691
0.55	0.26063	2.00	0.10908
0.60	0.26283	3.00	0.02680
0.65	0.26414	4.00	0.00517
0.69	0.26456	5.00	0.00090

The next regression model that we want to consider is given by the following set of equations:

$$Y_{1j} = \alpha_{11}e^{-\lambda_1 x_j} + \alpha_{12}e^{-\lambda_2 x_j} + \varepsilon_{1j} = E(Y_{1j}) + \varepsilon_{1j}$$

$$Y_{2j} = \alpha_{21}e^{-\lambda_1 x_j} + \alpha_{22}e^{-\lambda_2 x_j} + \varepsilon_{2j} = E(Y_{2j}) + \varepsilon_{2j} \qquad (7.20)$$

for $j = 0,1,2,\ldots,3M-1$. We also make the additional assumption that the vectors $\begin{pmatrix}\varepsilon_{1j}\\ \varepsilon_{2j}\end{pmatrix}$ are independent with each vector having a bivariate normal distribution with mean vector zero and covariance matrix $\Sigma = \begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{22}\end{pmatrix}$. With these assumptions the likelihood function is given by

$$L = \prod_{j=o}^{3M-1} \frac{1}{(2\pi)|\Sigma|^{\frac{1}{2}}} \exp\{-\tfrac{1}{2}[Y_{*j}-E(Y_{*j})]^T\Sigma^{-1}[Y_{*j}-E(Y_{*j})]\}$$

$$= \frac{1}{(2\pi)^{3M}|\Sigma|^{\frac{3M}{2}}} \exp\{-\frac{1}{2}\sum_{j=o}^{3M-1}[Y_{*j}-E(Y_{*j})]^T\Sigma^{-1}[Y_{*j}-E(Y_{*j})]\} \qquad (7.21)$$

where $Y_{*j} - E(Y_{*j}) = (Y_{1j}-E(Y_{1j}),\ Y_{2j}-E(Y_{2j}))^T$. Using equation (5.45) we note that $E\left(\frac{\partial \ln L}{\partial\lambda}\right)\left(\frac{\partial \ln L}{\partial\lambda}\right)^T = \sum_{j=o}^{3M-1} D_j\Sigma^{-1}D_j^T$ where

$$D_j = \begin{pmatrix} \dfrac{\partial E(Y_{1j})}{\partial \lambda_1} & \dfrac{\partial E(Y_{2j})}{\partial \lambda_1} \\ \\ \dfrac{\partial E(Y_{1j})}{\partial \lambda_2} & \dfrac{\partial E(Y_{2j})}{\partial \lambda_2} \end{pmatrix} = \begin{pmatrix} -\alpha_{11}x_j e^{-\lambda_1 x_j} & -\alpha_{21}x_j e^{-\lambda_1 x_j} \\ \\ -\alpha_{12}x_j e^{-\lambda_2 x_j} & -\alpha_{22}x_j e^{-\lambda_2 x_j} \end{pmatrix} . \tag{7.22}$$

Setting $x_j = hj$ and $J = hM$, we determine that

$$\sum_{j=o}^{3M-1} D_j \Sigma^{-1} D_j^T = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \text{ where}$$

$$A_{11} = (\sigma^{11}\alpha_{11}^2 + 2\sigma^{12}\alpha_{11}\alpha_{21} + \sigma^{22}\alpha_{21}^2) \sum_{j=o}^{3M-1} x_j^2 \, e^{-2\lambda_1 x_j}$$

$$= (\sigma^{11}\alpha_{11}^2 + 2\sigma^{12}\alpha_{11}\alpha_{21} + \sigma^{22}\alpha_{21}^2) \sum_{j=o}^{3M-1} \left(\frac{J}{M}\, j\right)^2 e^{-2\lambda_1 (\frac{J}{M} j)} \quad ; \tag{7.23a}$$

$$A_{12} = A_{21} = (\sigma^{11}\alpha_{11}\alpha_{12} + \sigma^{12}\alpha_{11}\alpha_{22} + \sigma^{12}\alpha_{12}\alpha_{21} + \sigma^{22}\alpha_{21}\alpha_{22})$$

$$\sum_{j=o}^{3M-1} x_j^2 e^{-(\lambda_1+\lambda_2)x_j}$$

$$= (\sigma^{11}\alpha_{11}\alpha_{12} + \sigma^{12}\alpha_{11}\alpha_{22} + \sigma^{12}\alpha_{12}\alpha_{21} + \sigma^{22}\alpha_{21}\alpha_{22})$$

$$\sum_{j=o}^{3M-1} \left(\frac{J}{M}\, j\right)^2 e^{-(\lambda_1+\lambda_2)(\frac{J}{M} j)} \quad ; \tag{7.23b}$$

$$A_{22} = (\sigma^{11}\alpha_{12}^2+2\sigma^{12}\alpha_{12}\alpha_{22}+\sigma^{22}\alpha_{22}^2) \sum_{j=o}^{3M-1} x_j^2 \, e^{-2\lambda_2 x_j}$$

$$= (\sigma^{11}\alpha_{12}^2 +2\sigma^{12}\alpha_{12}\alpha_{22}+\sigma^{22}\alpha_{22}^2) \sum_{j=o}^{3M-1} \left(\frac{J}{M}\, j\right)^2 e^{-2\lambda_2 (\frac{J}{M} j)} \quad ; \qquad (7.23c)$$

and

$$\Sigma^{-1} = \begin{pmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{12} & \sigma^{22} \end{pmatrix} .$$

Next we want to determine the expression for the asymptotic covariance matrix of the vector $\begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix}$, which is the vector of generalized partial totals estimators of the vector $\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$.

Without loss of generality let us assume that $\lambda_1 < \lambda_2$. Then the estimators of λ_1 and λ_2 are given by the following expressions:

$$\hat{\lambda}_k = - \frac{1}{hM} \ln w_k = - \frac{1}{J} \ln w_k \ , \qquad (7.24)$$

where

$$w_k = \frac{1}{2}\left\{L_1+(-1)^{k-1} \sqrt{L_1^2-4L_2}\ \right\}, \qquad (7.25)$$

for k = 1,2. That is, w_1 and w_2 are the roots of the quadratic equation

$$w^2 - L_1 w + L_2 = 0$$

where L_1 and L_2 are the estimators of the elementary symmetric functions of $e^{-\lambda_1 hM} = e^{-\lambda_1 J}$ and $e^{-\lambda_2 hM} = e^{-\lambda_2 J}$ found by solving the equations

$$\bar{S}_{12}L_1 - \bar{S}_{11}L_2 = \bar{S}_{13}$$

$$\bar{S}_{22}L_1 - \bar{S}_{21}L_2 = \bar{S}_{23}\ , \tag{7.26}$$

where $\bar{S}_{iq} = \frac{1}{M} \sum_{j=(q-1)M}^{qM-1} y_{ij}$ for $i = 1,2$ and $q = 1,2,3$. From Theorem 5.5 we note that the asymptotic covariance matrix of $\begin{matrix}\hat{\lambda}_1 \\ \hat{\lambda}_2\end{matrix}$ is given by

$$\frac{1}{M} F(\Sigma \otimes I)F^T \tag{7.27}$$

where a typical $(k, i^2+q-1)^{th}$ term of F is given by

$$\left.\frac{\partial \hat{\lambda}_k}{\partial \bar{S}_{iq}}\right|_{\bar{S}_{**}=\Psi_{**}}$$

for $i, k = 1,2$ and $q = 1,2,3$; $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$; I is a 3×3

identity matrix; $S_{**} = (\bar{S}_{11},\bar{S}_{12},\bar{S}_{13},\bar{S}_{21},\bar{S}_{22},\bar{S}_{23})^T$ and $\Psi_{**} = (\Psi_{11},\Psi_{12},\Psi_{13},\Psi_{21},\Psi_{22},\Psi_{23})^T$ where the Ψ_{iq}'s have been defined earlier in equation (5.23a). It can be shown that

$\frac{1}{M} F(\Sigma \otimes I)F^T = \frac{1}{M}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ where

$$B_{11} = \sigma_{11}\left[\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{11}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2 + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{12}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2 + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{13}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2\right]$$

$$+ \sigma_{22}\left[\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{21}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2 + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{22}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2 + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{23}}\ \bar{S}_{**}=\Psi_{**}\right)^2\right]$$

$$+ 2\sigma_{12}\left[\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{11}}\right)\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{21}}\right)\bigg|_{\bar{S}_{**}=\Psi_{**}} + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{12}}\right)\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{22}}\right)\bigg|_{\bar{S}_{**}=\Psi_{**}}\right.$$

$$\left. + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{13}}\right)\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{23}}\right)\bigg|_{\bar{S}_{**}=\Psi_{**}}\right]; \tag{7.28a}$$

$$B_{12} = B_{21} = \sigma_{11}\Bigg[\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{11}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{11}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}} + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{12}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{12}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}}$$

$$+ \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{13}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{13}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}}\Bigg]$$

$$+ \sigma_{22}\Bigg[\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{21}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{21}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}} + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{22}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{22}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}}$$

$$+\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{23}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{23}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}}\Bigg]$$

$$+ \sigma_{12}\Bigg[\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{21}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{11}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}} + \left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{22}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{12}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}}$$

$$+\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{23}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{13}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}} +\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{11}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{21}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}}$$

$$+\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{12}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{22}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}} +\left(\frac{\partial\hat{\lambda}_1}{\partial\bar{S}_{13}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{23}}\right)\Bigg|_{\bar{S}_{**}=\Psi_{**}}\Bigg];$$

(7.28b)

and

$$B_{22} = \sigma_{11}\left[\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{11}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2 + \left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{12}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2 + \left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{13}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2\right]$$

$$+ \sigma_{22}\left[\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{21}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2 + \left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{22}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2 + \left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{23}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)^2\right]$$

$$+ 2\sigma_{12}\left[\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{11}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{21}}\right)\bigg|_{\bar{S}_{**}=\Psi_{**}} + \left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{12}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{22}}\right)\bigg|_{\bar{S}_{**}=\Psi_{**}}\right.$$

$$\left. + \left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{13}}\right)\left(\frac{\partial\hat{\lambda}_2}{\partial\bar{S}_{23}}\right)\bigg|_{\bar{S}_{**}=\Psi_{**}}\right] . \tag{7.28c}$$

At the point Ψ_{**} we can show that $w_1 = e^{-\lambda_1 J}$ and $w_2 = e^{-\lambda_2 J}$. Therefore, from (7.24) it can be seen that

$$\frac{\partial\hat{\lambda}_k}{\partial\bar{S}_{iq}}\bigg|_{\bar{S}_{**}=\Psi_{**}} = -\frac{1}{Je^{-\lambda_k J}}\left(\frac{\partial w_k}{\partial\bar{S}_{iq}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right), \tag{7.29}$$

for $k = 1,2$. In addition we can show that

$$\left(\frac{\partial w_k}{\partial \bar{S}_{1q}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right) = \left\{(-1)^{k-1} e^{-\lambda_k J}\left(\frac{\partial L_1}{\partial \bar{S}_{iq}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)\right.$$

$$\left. + (-1)^k\left(\frac{\partial L_2}{\partial \bar{S}_{iq}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)\right\} \Big/ (e^{-\lambda_1 J} - e^{-\lambda_2 J}) \tag{7.30}$$

for $k = 1,2$. We now see that we will need the expressions for

$\left(\frac{\partial L_k}{\partial \bar{S}_{iq}}\bigg|_{\bar{S}_{**}=\Psi_{**}}\right)$ in order to determine the elements of the matrix

$\frac{1}{M}F(\Sigma \otimes I)F^T$, and we have listed these expressions in Table 7.8.

It can easily be seen that each element of the matrix $\frac{1}{M}F(\Sigma \otimes I)F^T$ has a common factor $\frac{1}{J^2 M}$. We factor this term out of this matrix and multiply each element of the matrix $\sum_{j=o}^{3M-1} D_j \Sigma^{-1} D_j^T$ by it. This means that the only elements of our expression for the efficiency involving M are terms of the form $\sum_{j=o}^{3M-1} x_j^2 e^{-\lambda x_j} \frac{1}{J^2 M}$. Using our previous results and the definition of a definite integral, we can show the following:

TABLE 7.8--Expressions for quantities appearing in equation (7.30)

$$\left.\frac{\partial L_1}{\partial \bar{S}_{11}}\right|_{\bar{S}_{**}=\Psi_{**}} = -e^{-\lambda_1 J-\lambda_2 J} C_{11}/E_1 \qquad \left.\frac{\partial L_2}{\partial \bar{S}_{11}}\right|_{\bar{S}_{**}=\Psi_{**}} = -e^{-\lambda_1 J-\lambda_2 J} C_{21}/E_1$$

$$\left.\frac{\partial L_1}{\partial \bar{S}_{12}}\right|_{\bar{S}_{**}=\Psi_{**}} = (e^{-\lambda_1 J}+e^{-\lambda_2 J})C_{11}/E_1 \qquad \left.\frac{\partial L_2}{\partial \bar{S}_{12}}\right|_{\bar{S}_{**}=\Psi_{**}} = (e^{-\lambda_1 J}+e^{-\lambda_2 J})C_{21}/E_1$$

$$\left.\frac{\partial L_1}{\partial \bar{S}_{13}}\right|_{\bar{S}_{**}=\Psi_{**}} = -C_{11}/E_1 \qquad \left.\frac{\partial L_2}{\partial \bar{S}_{13}}\right|_{\bar{S}_{**}=\Psi_{**}} = -C_{21}/E_1$$

$$\left.\frac{\partial L_1}{\partial \bar{S}_{21}}\right|_{\bar{S}_{**}=\Psi_{**}} = e^{-\lambda_1 J-\lambda_2 J} C_{12}/E_1 \qquad \left.\frac{\partial L_2}{\partial \bar{S}_{21}}\right|_{\bar{S}_{**}=\Psi_{**}} = e^{-\lambda_1 J-\lambda_2 J} C_{22}/E_1$$

$$\left.\frac{\partial L_1}{\partial \bar{S}_{22}}\right|_{\bar{S}_{**}=\Psi_{**}} = -(e^{-\lambda_1 J}+e^{-\lambda_2 J})C_{12}/E_1 \qquad \left.\frac{\partial L_2}{\partial \bar{S}_{22}}\right|_{\bar{S}_{**}=\Psi_{**}} = -(e^{-\lambda_1 J}+e^{-\lambda_2 J})C_{22}/E_1$$

$$\left.\frac{\partial L_1}{\partial \bar{S}_{23}}\right|_{\bar{S}_{**}=\Psi_{**}} = C_{12}/E_1 \qquad \left.\frac{\partial L_2}{\partial \bar{S}_{23}}\right|_{\bar{S}_{**}=\Psi_{**}} = C_{22}/E_1$$

$$C_{11} = \alpha_{21} J\lambda_2(1-e^{-\lambda_1 J}) + \alpha_{22} J\lambda_1(1-e^{-\lambda_2 J})$$

$$C_{12} = \alpha_{11} J\lambda_2(1-e^{-\lambda_1 J}) + \alpha_{12} J\lambda_1(1-e^{-\lambda_2 J})$$

$$C_{21} = \alpha_{21} J\lambda_2 e^{-\lambda_1 J}(1-e^{-\lambda_1 J}) + \alpha_{22} J\lambda_1 e^{-\lambda_2 J}(1-e^{-\lambda_2 J})$$

$$C_{22} = \alpha_{11} J\lambda_2 e^{-\lambda_1 J}(1-e^{-\lambda_1 J}) + \alpha_{12} J\lambda_1 e^{-\lambda_2 J}(1-e^{-\lambda_2 J})$$

$$E_1 = (1-e^{-\lambda_1 J})(1-e^{-\lambda_2 J})(e^{-\lambda_1 J}-e^{-\lambda_2 J})(\alpha_{12}\alpha_{21}-\alpha_{11}\alpha_{22})$$

$$\lim_{M\to\infty} \sum_{j=o}^{3M-1} x_j^2 e^{-\lambda x_j} \frac{1}{J^2 M} = \frac{1}{J^3} \lim_{M\to\infty} \sum_{j=o}^{3M-1} \left(\frac{J}{M} j\right)^2 e^{-\lambda(\frac{J}{M}j)} \frac{J}{M} = \frac{1}{J^3} \int_o^{3J} t^2 e^{-\lambda t} dt$$

$$= \left(2-2e^{-3\lambda J} - 6\lambda J e^{-3\lambda J} - 9\lambda^2 J^2 e^{-3\lambda J}\right) \Big/ \lambda^3 J^3. \tag{7.31}$$

After the determination of the form of each of the elements in the expression for the asymptotic efficiency, ν, of $\begin{matrix}\hat{\lambda}_1 \\ \hat{\lambda}_2\end{matrix}$, we let $\Sigma = \begin{pmatrix} 0.001 & 0.0009 \\ 0.0009 & 0.001 \end{pmatrix}$, $\alpha_{12} = 1 - \alpha_{11}$, $\alpha_{22} = 1 - \alpha_{21}$, and evaluated the asymptotic efficiency for the following ranges of the parameters: $0.10 \leq \alpha_{11} \leq 0.80$, $0.25 \leq \alpha_2 \leq 0.90$, $0.004 \leq J\lambda_1 \leq 1.9$, and $1.0 \leq J\lambda_2 \leq 4.0$. For the particular models considered, we found that in every instance $\nu < 0.002$. However, it was possible to note during these calculations that ν did achieve its smallest values when $|J\lambda_1 - J\lambda_2|$, or equivalently $|\lambda_1 - \lambda_2|$, was relatively small.

7.3.2 Generalized Spearman estimation procedure

In this section we evaluate the asymptotic efficiency of the generalized Spearman estimators of the exponential parameters for some particular regression models. The expression for the asymptotic efficiency of the estimators of the exponential parameters is given in Section 6.4 and defined in equation (6.50).

The first model that we consider is the case when $n = 1$ and the random variable Y_{1j} represents the proportion of "successes" in n^* independent binomial trials. Therefore we may write $Y_{1j} = \frac{r_{1j}}{n^*}$ and we assume that

$$E(Y_{1j}) = \exp(-\lambda_1 e^{z_j}) \tag{7.32}$$

where the values of z_j have been specified in Section 6.2 for $j = -M', -M'+1,\ldots,0,1,\ldots,M''$. This is the same model and experimental situation as considered by Johnson and Brown [1961], and also this model is among the general class of models given in equation (6.5c) with $n = 1$ and $\alpha_{11} = 1$. Using the same assumptions as stated in Section 6.4, we may write the likelihood function as

$$L = \prod_{j=-M'}^{M''} \binom{n^*}{r_{1j}} p_{1j}^{r_{1j}} (1-p_{1j})^{n^*-r_{1j}} \tag{7.33}$$

where $p_{1j} = E(Y_{1j})$. It can then be shown that

$$\left(\frac{\partial \ln L}{\partial \lambda_1}\right)^2 = \sum_{j=-M'}^{M''} (p_{1j}^{(1)})^2 \frac{(r_{1j}-n^*p_{1j})^2}{p_{1j}^2(1-p_{1j})^2}$$

$$+ \sum_{\substack{j,j'=-M' \\ j\neq j'}}^{M''} (p_{1j}^{(1)})(p_{1j'}^{(1)}) \frac{(r_{1j}-n^*p_{1j})(r_{1j'}-n^*p_{1j'})}{p_{1j}(1-p_{1j})p_{1j'}(1-p_{1j'})}$$

where $p_{1j}^{(1)} = \partial p_{1j}/\partial\lambda_1$. Hence we now have the following expression for $E\left(\frac{\partial \ln L}{\partial\lambda_1}\right)^2$:

$$E\left(\frac{\partial \ln L}{\partial\lambda_1}\right)^2 = \sum_{j=-M'}^{M''} (p_{1j}^{(1)})^2 \frac{n^*}{p_{1j}(1-p_{1j})} . \tag{7.34}$$

From the results of Section 6.2, we note that the estimator of λ_1 is given by

$$\hat{\lambda}_1 = e^{-r} e^{-\hat{\mu}_1^{(1)}} \tag{7.35}$$

where

$$\hat{\mu}_1^{(1)} = \sum_{j=-M'}^{M''-1} \left(\frac{z_j+z_{j+1}}{2}\right) \Delta y_{1j} \tag{7.36}$$

and y_{1j} represents the observed value of Y_{1j} at the point z_j and $\Delta y_{1j} = y_{1,j+1} - y_{1j}$. From Theorem 6.7 we note that the asymptotic variance of $\hat{\lambda}_1$ is given by

$$F' \left\{ d^2 \sum_{j=-M'+1}^{M''-2} \frac{p_{1j}(1-p_{1j})}{n^*} \right\} F'^T \tag{7.37}$$

where

$$F' = \left(\frac{\partial \hat{\lambda}_1}{\partial \hat{\mu}_1^{(1)}} \Bigg|_{\hat{\mu}_1^{(1)} = \mu_1^{(1)}} \right) = - e^{-\gamma} e^{-\mu_1^{(1)}} = - \lambda_1 . \tag{7.38}$$

Hence the asymptotic variance of $\hat{\lambda}_1$ is given by

$$\frac{\lambda_1^2 d^2}{n^*} \sum_{j=-M'+1}^{M''-2} p_{1j} (1-p_{1j}). \tag{7.39}$$

Combining the results from equations (7.34) and (7.39), the reciprocal of the asymptotic efficiency of $\hat{\lambda}_1$ is given by

$$\frac{1}{\nu} = \lim_{M' \to \infty} \left(n^* \sum_{j=-M'}^{M''} \frac{(p_{1j}^{(1)})^2}{p_{1j}(1-p_{1j})} \right) \left(\frac{\lambda_1^2 d^2}{n^*} \sum_{j=-M'+1}^{M''-2} p_{1j}(1-p_{1j}) \right)$$

$$= \lambda_1^2 \lim_{M' \to \infty} \left(d \sum_{j=-M'}^{M''} \frac{(p_{1j}^{(1)})^2}{p_{1j}(1-p_{1j})} \right) \left(d \sum_{j=-M'+1}^{M''-2} p_{1j}(1-p_{1j}) \right)$$

$$= \lambda_1^2 \left[\int_{-\infty}^{\infty} \frac{(p_1^{(1)}(z))^2}{p_1(z)(1-p_1(z))} \, dz \right] \left[\int_{-\infty}^{\infty} p_1(z)[1-p_1(z)] \, dz \right] \tag{7.40a}$$

where $p_1^{(1)}(z) = \frac{\partial p_1(z)}{\partial \lambda_1}$ and $p_1(z) = E(Y_1(z)) = \exp(-\lambda_1 e^z)$. If we make the substitution $t = \exp(-\lambda_1 e^z)$ in the last two integrals of equation (7.40a) then

$$\frac{1}{\nu} = \left[-\int_0^1 \frac{\ln t}{(1-t)} \, dt \right] \left[\int_0^1 \frac{(t-1)}{\ln t} \, dt \right] = \frac{\pi^2 \ln 2}{6} \tag{7.40b}$$

or

$$\nu = \frac{6}{\pi^2 \ln 2} = 0.87705. \tag{7.40c}$$

It is interesting to note the high asymptotic efficiency for this model, which verifies the result quoted by Johnson and Brown [1961], and also to note that ν is independent of n^*.

As an extension of the results just presented, next we will consider the case where $Y_{1j} = r_{1j}/n^*$, $Y_{2j} = r_{2j}/n^*$, and r_{1j} and r_{2j} are random variables having a multinomial distribution. For this particular case we assume that

$$E(Y_{1j}) = \alpha_{11} \exp\,(-\lambda_1 e^{z_j}) + (1-\alpha_{11}) \exp\,(-\lambda_2 e^{z_j}) = p_{1j}$$

and

$$E(Y_{2j}) = 1-\alpha_{21} \exp\,(-\lambda_1 e^{z_j}) - (1-\alpha_{21}) \exp\,(-\lambda_2 e^{z_j}) = p_{2j}, \tag{7.41}$$

for $j = -M', -M'+1, \ldots, 0, 1, \ldots, M''$, which is the example given by equation (6.5c) with $n = 2$. Using the assumptions of Section 6.4, the likelihood function for this example is given by

$$L = \prod_{j=-M'}^{M''} \frac{n^*!}{r_{1j}!\,r_{2j}!\,(n^*-r_{1j}-r_{2j})!}\, p_{1j}^{r_{1j}}\, p_{2j}^{r_{2j}}\, (1-p_{1j}-p_{2j})^{n^*-r_{1j}-r_{2j}}. \tag{7.42}$$

After extensive algebra we find

$$\frac{\partial \ln L}{\partial \lambda_k} = \sum_{j=-M'}^{M''} \left\{ p_{1j}^{(k)} \left[\frac{(1-p_{2j})(r_{1j}-n^{*}p_{1j})+p_{1j}(r_{2j}-n^{*}p_{2j})}{p_{1j}(1-p_{1j}-p_{2j})} \right] + p_{2j}^{(k)} \left[\frac{(1-p_{1j})(r_{2j}-n^{*}p_{2j})+p_{2j}(r_{1j}-n^{*}p_{1j})}{p_{2j}(1-p_{1j}-p_{2j})} \right] \right\} \tag{7.43}$$

and

$$\left(\frac{\partial \ln L}{\partial \lambda_1}\right)\left(\frac{\partial \ln L}{\partial \lambda_2}\right) = \sum_{j=-M'}^{M''} \left\{ p_{1j}^{(1)} \left[\frac{(1-p_{2j})(r_{1j}-n^{*}p_{1j})+p_{1j}(r_{2j}-n^{*}p_{2j})}{p_{1j}(1-p_{1j}-p_{2j})} \right] + p_{2j}^{(1)} \left[\frac{(1-p_{1j})(r_{2j}-n^{*}p_{2j})+p_{2j}(r_{1j}-n^{*}p_{1j})}{p_{2j}(1-p_{1j}-p_{2j})} \right] \right\}$$

$$\left\{ p_{1j}^{(2)} \left[\frac{(1-p_{2j})(r_{1j}-n^{*}p_{1j})+p_{1j}(r_{2j}-n^{*}p_{2j})}{p_{1j}(1-p_{1j}-p_{2j})} \right] + p_{2j}^{(2)} \left[\frac{(1-p_{1j})(r_{2j}-n^{*}p_{2j})+p_{2j}(r_{1j}-n^{*}p_{1j})}{p_{2j}(1-p_{1j}-p_{2j})} \right] \right\}$$

$$+\sum_{\substack{j,j'=-M' \\ j\neq j'}}^{M''}\left\{p_{1j}^{(1)}\left[\frac{(1-p_{2j})(r_{1j}-n^{*}p_{1j})+p_{1j}(r_{2j}-n^{*}p_{2j})}{p_{1j}(1-p_{1j}-p_{2j})}\right]\right.$$

$$\left.+\,p_{2j}^{(1)}\left[\frac{(1-p_{1j})(r_{2j}-n^{*}p_{2j})+p_{2j}(r_{1j}-n^{*}p_{1j})}{p_{2j}(1-p_{1j}-p_{2j})}\right]\right\}$$

$$\left\{p_{1j'}^{(2)}\left[\frac{(1-p_{2j'})(r_{1j'}-n^{*}p_{1j'})+p_{1j'}(r_{2j'}-n^{*}p_{2j'})}{p_{1j'}(1-p_{1j'}-p_{2j'})}\right]\right.$$

$$\left.+\,p_{2j'}^{(2)}\left[\frac{(1-p_{1j'})(r_{2j'}-n^{*}p_{2j'})+p_{2j'}(r_{1j'}-n^{*}p_{1j'})}{p_{2j'}(1-p_{1j'}-p_{2j'})}\right]\right\} \tag{7.44}$$

where $p_{ij}^{(k)} = \partial p_{ij}/\partial\lambda_k$. From the assumptions that we have made about the random variables r_{ij}, we know that Var $(r_{ij})=n^{*}p_{ij}(1-p_{ij})$ and Cov $(r_{1j},r_{2j})=-n^{*}p_{1j}p_{2j}$. Hence we may conclude that

$$E\left(\frac{\partial \ln L}{\partial\lambda_k}\right)^2 = n^{*}\sum_{j=-M'}^{M''}\left\{\frac{1}{(1-p_{1j}-p_{2j})}\right\}$$

$$\left\{\frac{(p_{1j}^{(k)})^2(1-p_{2j})}{p_{1j}} + \frac{(p_{2j}^{(k)})^2(1-p_{1j})}{p_{2j}} + 2p_{1j}^{(k)}\,p_{2j}^{(k)}\right\} \tag{7.45}$$

for k=1,2, and

$$E\left(\frac{\partial \ln L}{\partial \lambda_1}\right)\left(\frac{\partial \ln L}{\partial \lambda_2}\right) = n^* \sum_{j=-M'}^{M''} \left\{\frac{1}{(1-p_{1j}-p_{2j})}\right\}$$

$$\left\{\frac{p_{1j}^{(1)} p_{1j}^{(2)} (1-p_{2j})}{p_{1j}} + p_{2j}^{(1)} p_{2j}^{(2)} \frac{(1-p_{1j})}{p_{2j}} + p_{1j}^{(1)} p_{2j}^{(2)} + p_{1j}^{(2)} p_{2j}^{(1)}\right\}. \qquad (7.46)$$

Next we want to determine the expressions for the asymptotic covariance matrix of the generalized Spearman estimators of the vector $\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}$. Let $p'_{2j} = 1-p_{2j}$; then both p_{1j} and p'_{2j} fall under the classification of Case 1 of Chapter 6. Without loss of generality let us assume that $\lambda_1 > \lambda_2$. Then the estimators of λ_1 and λ_2 are given by the following expressions:

$$\hat{\lambda}_k = e^{w_k}, \qquad (7.47)$$

where

$$w_k = \frac{1}{2}\left\{L_1^{(C)} + (-1)^{k-1}\sqrt{L_1^{(C)2} - 4L_2^{(C)}}\right\}, \qquad (7.48)$$

for $k = 1,2$.

That is w_1 and w_2 are the two roots of the equation

$$w^2 - L_1 w + L_2 = 0 , \qquad (7.49)$$

where $L_1^{(C)}$ and $L_2^{(C)}$ are the estimators of the elementary symmetric functions of $\ln\lambda_1$ and $\ln\lambda_2$ found by solving the equations

$$\hat{K}_{11} L_1^{(C)} - L_2^{(C)} = -\hat{K}_{12}$$

$$\hat{K}_{21} L_1^{(C)} - L_2^{(C)} = -\hat{K}_{22} \qquad (7.50)$$

where $\hat{K}_{11} = \hat{\mu}_1^{(1)} + I_1$; $\hat{K}_{12} = \hat{\mu}_1^{(2)} + I_2 - 2I_1\hat{K}_{11}$; $\hat{K}_{21} = \hat{\mu}_2^{\prime(1)} + I_1$; and $\hat{K}_{22} = \hat{\mu}_2^{\prime(2)} + I_2 - 2I_1\hat{K}_{21}$. The expressions for $\mu_1^{(1)}$ and $\mu_1^{(2)}$ are given in Theorem 6.1 where $p_1(z) = E(Y_1(z) = \alpha_{11}\exp(-\lambda_1 e^z) + (1-\alpha_{11})\exp(-\lambda_2 e^z)$, and the quantities $\mu_2^{\prime(1)}$ and $\mu_2^{\prime(2)}$ are defined in the same theorem using $p_2^\prime(z) = 1-p_2(z) = 1-E(Y_2(z)) = \alpha_{21}\exp(-\lambda_1 e^z) + (1-\alpha_{21})\exp(-\lambda_2 e^z)$ in the place of $p_2(z)$. The estimators of $\mu_1^{(1)}$, $\mu_1^{(2)}$, $\mu_2^{\prime(1)}$, and $\mu_2^{\prime(2)}$ are defined in equation (6.35a), where we use the observations $y_{2j}^\prime = 1 - y_{2j}$ in the place of y_{2j} in $\hat{\mu}_2^{\prime(1)}$ and $\hat{\mu}_2^{\prime(2)}$. From Theorem 6.7 we note that the asymptotic covariance matrix of $\begin{pmatrix}\hat{\lambda}_1\\ \hat{\lambda}_2\end{pmatrix}$ is given by

$$F^\prime \, \Omega \, F^{\prime T} \tag{7.51}$$

where $F^\prime$ has been defined in connection with equation (6.48) and

$$\Omega = \frac{d^2}{n^*} \sum_{j=-M^\prime+1}^{M^{\prime\prime}-2} \begin{pmatrix} p_{1j}(1-p_{1j}) & 2z_j p_{1j}(1-p_{1j}) & p_{1j}p_{2j} & 2z_j p_{1j}p_{2j} \\ 2z_j p_{1j}(1-p_{1j}) & 4z_j^2 p_{1j}(1-p_{1j}) & 2z_j p_{1j}p_{2j} & 4z_j^2 p_{1j}p_{2j} \\ p_{1j}p_{2j} & 2z_j p_{1j}p_{2j} & p_{2j}(1-p_{2j}) & 2z_j p_{2j}(1-p_{2j}) \\ 2z_j p_{1j}p_{2j} & 4z_j^2 p_{1j}p_{2j} & 2z_j p_{2j}(1-p_{2j}) & 4z_j^2 p_{2j}(1-p_{2j}) \end{pmatrix}. \tag{7.53}$$

It can be seen that the factor $\frac{1}{n^*}$ in the asymptotic covariance matrix cancels with the factor n^* in the matrix $E\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T$. Hence our expression for the asymptotic efficiency, ν, of $\begin{pmatrix}\hat{\lambda}_1\\ \hat{\lambda}_2\end{pmatrix}$ will be independent of n^*. We now factor out a d from the asymptotic covariance matrix and multiply this d times the matrix $E\left(\frac{\partial \ln L}{\partial \lambda}\right)\left(\frac{\partial \ln L}{\partial \lambda}\right)^T$. After doing this we find the reciprocal of the asymptotic efficiency to be given by

$$\frac{1}{\nu} = \lim_{M' \to \infty} \left| dE \; \frac{\partial \ln L}{\partial \lambda} \; \frac{\partial \ln L}{\partial \lambda}^T \right| \left| \frac{1}{d} F' \; \Omega \; F'^T \right|$$

$$= \begin{vmatrix} AA_{11} & AA_{12} \\ AA_{21} & AA_{22} \end{vmatrix} \left| F' \begin{pmatrix} BB_{11} & BB_{12} & BB_{13} & BB_{14} \\ BB_{21} & BB_{22} & BB_{23} & BB_{24} \\ BB_{31} & BB_{32} & BB_{33} & BB_{34} \\ BB_{41} & BB_{42} & BB_{43} & BB_{44} \end{pmatrix} F'^T \right| , \quad (7.54)$$

where the expressions for the AA's and BB's are given by the following:

$$AA_{11} = \int_{-\infty}^{\infty} \{1/(1-p_1(z)-p_2(z))\}\{(p_1^{(1)}(z))^2 \; [(1-p_2(z)/p_1(z)]$$

$$+(p_2^{(1)}(z))^2[(1-p_1(z))/p_2(z)]+2p_1^{(1)}(z)p_2^{(1)}(z)\} \; dz \; ;$$

$$AA_{12}=AA_{21}=\int_{-\infty}^{\infty}\{1/(1-p_1(z)-p_2(z))\}\{(p_1^{(1)}(z)p_1^{(2)}(z))[(1-p_2(z))/p_1(z)]$$

$$+(p_1^{(1)}(z)p_2^{(2)}(z)+p_1^{(2)}(z)p_2^{(1)}(z))$$

$$+(p_2^{(1)}(z)p_2^{(2)}(z))[(1-p_1(z))/p_2(z)]\}\ dz;$$

$AA_{22}=AA_{11}$ with $p_1^{(1)}(z)$ replaced by $p_1^{(2)}(z)$ and $p_2^{(1)}(z)$ replaced by $p_2^{(2)}(z)$;

$$BB_{13}=BB_{31}=\int_{-\infty}^{\infty} p_1(z)p_2(z)dz;$$

$$BB_{11}=\int_{-\infty}^{\infty} p_1(z)(1-p_1(z))\ dz;$$

$$BB_{33}=\int_{\infty}^{\infty} p_2(z)(1-p_2(z))\ dz;$$

$$BB_{14}=BB_{41}=BB_{23}=BB_{32}=2\int_{-\infty}^{\infty} zp_1(z)p_2(z)dz;$$

$$BB_{12}=BB_{21}=2\int_{-\infty}^{\infty} zp_1(z)(1-p_1(z))dz;$$

$$BB_{34}=BB_{43}=2\int_{-\infty}^{\infty} zp_2(z)(1-p_2(z))\ dz;$$

$$BB_{24}=BB_{42}=4\int_{-\infty}^{\infty} z^2p_1(z)p_2(z)\ dz;$$

$$BB_{22}=4\int_{-\infty}^{\infty} z^2p_1(z)(1-p_1(z))\ dz;$$

and

$$BB_{44}=4\int_{-\infty}^{\infty} z^2p_2(z)(1-p_2(z))dz.$$

We have been able to demonstrate the existence of the integrals given by AA_{11}, AA_{12}, and AA_{22}, but we have been unable to evaluate these integrals in a closed form. Therefore we have resorted to numerical methods for the evaluation of these integrals. A detailed discussion of the quadrature method that we used for the evaluation of these integrals is given in the book by Ralston and Wilf ([1960], page 242-248).

In Table 7.9, we tabulate the values of the asymptotic efficiency of $\begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix}$ for various values of the parameters of the model (7.41). In addition to the values given in Table 7.9, we also attempted to calculate the value of ν for $\alpha_{11} = 0.70$; $\alpha_{21} = 0.35$; $\lambda_1 = 0.85$; and $\lambda_2 = 0.04$. Because of some problems with the numerical integration of the integrals for AA_{11}, AA_{12}, and AA_{22}, we are not certain of the number of significant figures for the value of ν, which came out to be approximately 0.90. From a visual examination of this table it can be seen that ν becomes very small as $(\lambda_1-\lambda_2)$ becomes small. This result is not inconsistent with what we would expect, since for those cases when $(\lambda_1-\lambda_2)$ is small our model could just as well be reduced to two single exponential equations. As $(\lambda_1-\lambda_2)$ increases, we note that ν does take on some moderate values. In addition, this table also shows, for the model being considered, that a similar relation exists between ν and $(\alpha_{11}-\alpha_{21})$.

Although we have not considered in this research the specific problem of constructing confidence ellipsoids, this could be another problem for consideration in future research. This general problem is related to the problem of measuring the nonlinearity of a

TABLE 7.9--Asymptotic efficiencies of the generalized Spearman estimates of the exponential parameters in (7.41)

λ_1	λ_2	α_{11}	α_{21}	ν
0.75	0.55	0.80	0.10	0.00016
0.95	0.65	0.80	0.10	0.00034
0.85	0.55	0.80	0.10	0.00058
0.95	0.55	0.80	0.10	0.00141
0.95	0.095	0.55	0.45	0.11725
0.90	0.080	0.55	0.45	0.15063
0.90	0.070	0.55	0.45	0.18792
0.875	0.065	0.55	0.45	0.20988
0.80	0.060	0.55	0.45	0.23140
0.85	0.060	0.55	0.45	0.23740
0.95	0.095	0.70	0.35	0.24239
0.90	0.080	0.70	0.35	0.31326
0.90	0.070	0.70	0.35	0.39507
0.875	0.065	0.70	0.35	0.44136
0.80	0.060	0.70	0.35	0.48052
0.85	0.060	0.70	0.35	0.49879

nonlinear model, which was considered by Beale [1960] and Guttman and Meeter [1965] and mentioned in Chapter 2 of this research. We could test, by using the measure of nonlinearity given in equation (2.17), to see if our models may reasonably be approximated by a linear model. If the measure of nonlinearity is small, then according to Beale, linear regression theory results may be used to construct approximate confidence regions.

7.4 Some remarks

Although it must be remembered that the conclusions we arrive at in this chapter are directly related to the particular models considered, it might be useful to someone desiring to apply one of the generalized estimation procedures developed in this research to briefly summarize some of the results of these examples.

1) From a visual examination of the fitted equations in Figures 7.3 through 7.15, there does not appear to be a great deal of difference between the equations fitted by the generalized partial totals procedure and the generalized least squares procedure, or the generalized Spearman estimation procedure and generalized least squares. All of the procedures appear to give reasonably good fits.

2) For the particular model considered in this chapter, the values of the asymptotic efficiency of the generalized partial totals estimators of the exponential parameters were very small. Even in the single exponential model the asymptotic efficiency was always less than 0.27. So this particular criterion does not indicate that

this estimation procedure has a great deal to offer when samples are taken for a large number of points. However, if the assumptions concerning the random variables and the spacing of the observations of our model are satisfied, then this generalized partial totals technique is easy to apply, compared to the generalized least squares or iterative maximum likelihood procedure. In any case, we can use these partial totals estimates as initial estimates for the generalized least squares procedure.

3) For the particular model considered, the values of the asymptotic efficiency of the generalized Spearman estimators of the exponential parameters did achieve some moderately high values. In addition, this technique has the advantage of being a simple technique like the partial totals procedure. Although it may be difficult to satisfy the assumptions concerning the spacing of the observations and the random variables of a model under consideration, we have already indicated in previous chapters how these assumptions may be satisfied in particular experimental situations, e.g. in numerous biological serial dilution experiments the observations are taken at exponentially spaced values of the independent variable. Just as we mentioned in connection with the generalized partial totals estimators, we may always use the Spearman estimators as initial estimators for the generalized least squares procedure. However, because of its simplicity and relatively high asymptotic efficiency for a model like (7.41), this method would be preferred

to generalized least squares estimation for a model like (7.41) if it is reasonable to assume that the coefficients as well as the exponential parameters differ widely.

VIII. SUMMARY AND CONCLUSIONS

In this research we have generalized three nonlinear estimation procedures so that we can apply them simultaneously to a multiple equation regression model. These three generalized procedures have been designated by the following names:

(1) generalized least squares estimation procedure;

(2) generalized partial totals estimation procedure; and

(3) generalized Spearman estimation procedure. We have shown how the generalized least squares procedure may be applied to the estimation of the parameters in the regression model given by (1.2). Also we have shown how the other two generalized procedures may be applied to the estimation of the parameters for some of the members of the class of regression models specified by (1.1) which arise in the analysis of compartmental models. In order to give some motivation to the consideration of the class of regression models given by (1.1), we have devoted the third chapter to a discussion of tracer experiments and various compartmental models.

In addition to the development and generalization of these estimation procedures, we have also considered some of the asymptotic properties of our estimators. We have demonstrated the consistency of our estimators, and for the generalized partial totals and Spearman estimation procedures we have derived the limiting distribution of the estimators of the nonlinear parameters. We have also defined a measure of efficiency of the estimators of the nonlinear parameters for the generalized partial totals and generalized Spearman estimation procedures.

In the seventh chapter, we have applied the three generalized estimation procedures to some sets of data from particular regression models. For those sets of data from which we calculated our generalized partial totals and generalized Spearman estimates, we also calculated the generalized least squares estimates of the parameters of our model. We have also displayed some graphs of the original data with the fitted regression equations. which may serve as a visual comparison of the various techniques for the particular models considered. In this chapter we have also evaluated, for some particular regression models, the expressions for the asymptotic efficiency of the estimators of the exponential parameters found by the generalized partial totals and generalized Spearman estimation procedures. This allows for another comparison of the various estimation techniques.

Although the limited empirical comparisons contained in the seventh chapter of this research do not allow a basis to judge the various generalized estimation procedures in general, there are some points that have been evident throughout this research. The generalized least squares procedure is applicable to a larger class of regression models than the two other generalized procedures, but the difficulty of applying this procedure and the problems of convergence may be more important than the advantage of wide applicability. The generalized partial totals procedure has been extended to the estimation of the parameters in a particular subclass of regression models given by (1.1) when the observations are equally spaced, and the estimates found by this procedure are much more easily obtained than the least squares estimates. The generalized Spearman estimation procedure also has the desirable characteristic of being easily applied to the estimation of the parameters in a particular subclass of the regression models given by (1.1), when the observations are equally spaced on a logarithmic scale, and, in addition, has the advantage that the asymptotic efficiency of the estimators of the exponential parameters achieves moderately high values. Therefore if a model, to which the generalized partial totals or generalized Spearman estimation procedures may be applied, is appropriate, then these procedures provide attractive alternatives to the complicated generalized least squares procedure.

REFERENCES

Aitken, A. C. [1934]. On least-squares and linear combination of observations. Proceedings of the Royal Society of Edinburgh 55, 42-48.

Aitken, A. C. [1949]. Determinants and Matrices. Oliver and Boyd, Edinburgh and London.

Anderson, T. W. [1958]. An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, Incorporated, New York.

Beale, E. M. L. [1960]. Confidence regions in nonlinear estimation. J. of the Royal Statist. Soc. B22, 41-76.

Beauchamp, J. J., and Cornell, R. G. [1966]. Simultaneous nonlinear estimation. Technometrics 8, 319-326.

Berman, M. [1961]. Application of differential equations to the study of the thyroid system. Proceedings of the Fourth Berkeley Symposium, Vol. IV. University of California Press, Berkeley and Los Angeles. 87-99.

Berman, M., and Schoenfeld, R. [1956]. Invariants in experimental data on linear kinetics and the formulation of models. J. of Applied Physics 27, 1361-1370.

Birnbaum, A. [1954]. Statistical methods for Poisson processes and exponential populations. J. of the Amer. Statist. Assoc. 49, 254-266.

Box, G. E. P., and Draper, N. R. [1965]. The Bayesian estimation of common parameters from several responses. Biometrika 52, 355-364.

Bross, Irwin [1954]. A confidence interval for a percentage increase. Biometrics 10, 245-250.

Cornell, R. G. [1956]. A new estimation procedure for linear combinations of exponentials. Oak Ridge National Laboratory Report ORNL - 2120, Oak Ridge, Tennessee.

Cornell, R. G. [1962]. A method for fitting linear combinations of exponentials. Biometrics 18, 104-113.

Cornell, R. G. [1965]. Spearman estimation for a simple exponential model. Biometrics 21, 858-864.

Cornfield, J., Steinfeld, J., and Greenhouse, S. W. [1960]. Models for the interpretation of experiments using tracer compounds. Biometrics 16, 212-234.

Cramér, H. [1946]. Mathematical Methods of Statistics. Princeton University Press, Princeton.

Cramér, H. [1937]. Random Variables and Probability Distributions. Cambridge University Press, Cambridge.

Curtiss, J. H. [1941]. On the distribution of the quotient of two chance variables. Ann. Math. Statist. 12, 409-421.

Donahue, J. D. [1964]. Products and quotients of random variables and their applications. Aerospace Research Laboratories Report ARL 64-115.

Fieller, E. C. [1932]. The distribution of the index in a normal bivariate population. Biometrika 24, 428-440.

Fisher, R. A. [1921]. On the mathematical foundations of theoretical statistics. Philosophical Transactions of the Royal Society A222, 309-368.

Galambos, J. T., and Cornell, R. G. [1962]. Mathematical models for the study of the metabolic pattern of sulfate. The Journal of Laboratory and Clinical Medicine 60, 53-63.

Geary, R. C. [1930]. The frequency distribution of the quotient of two normal variates. J. of the Royal Statist. Soc. A93, 442-446.

Gurland, J. [1948]. Inversion formulae for the distribution of ratios. Ann. Math. Statist. 19, 228-237.

Guttman, I., and Meeter. D. A. [1965]. On Beale's measures of non-linearity. Technometrics 7, 623-637.

Hartley, H. O. [1948]. The estimation of non-linear parameters by 'internal least squares'. Biometrika 35, 32-45.

Hartley, H. O. [1961]. The modified Gauss-Newton method for the fitting of non-linear regression functions by least squares. Technometrics 3, 269-280.

Hartley, H. O., and Booker, A. [1965]. Nonlinear least squares estimation. Ann. Math. Statist. 36, 638-650.

Hohn, F. E. [1964]. Elementary Matrix Algebra. The Macmillan Company, New York.

Johnson, E. A., and Brown, B. W., Jr. [1961]. The Spearman estimator for serial dilution assays. Biometrics 17, 79-88.

Kendall, M. G., and Stuart, A. [1961]. The Advanced Theory of Statistics, Vol. II. Hafner Publishing Company, New York.

Levenberg, K. [1944]. A method for the estimation of certain non-linear problems in least squares. Quarterly of Applied Mathematics 2, 164-168.

Lipton, S., and McGilchrist, C. A. [1964]. The derivation of methods for fitting exponential regression curves. Biometrika 51, 504-508.

Marquardt, D. W. [1963]. An algorithm for least-squares estimation of nonlinear parameters. J. for the Soc. for Industrial and Applied Mathematics 11, 431-441.

Merrill, A. S. [1928]. Frequency distribution of an index when both the components follows the normal law. Biometrika 20A, 53-63.

Pratt, J. W. [1959]. On a general concept of 'in probability'. Ann. Math. Statist. 30, 549-558.

Ralston, A., and Wilf, H. S. [1960]. Mathematical Methods for Digital Computers. John Wiley and Sons, Incorporated, New York.

Sheppard, C. W. [1962]. Basic Principles of the Tracer Method. John Wiley and Sons, Incorporated, New York.

Speckman, J. A., and Cornell, R. G. [1965]. Estimation for a one-parameter exponential model. J. of the Amer. Statist. Assoc. 60, 560-572.

Stevens, W. L. [1951]. Asymptotic regression. Biometrics 7, 247-267.

Sverdrup, E. [1952]. The limit distribution of a continuous function of random variables. Skankinavisk Aktuarietidskrift 35, 1-10.

Telser, L. G. [1964]. Iterative estimaticn of a set of linear regression equations. J. of the Amer. Statist. Assoc. 59, 845-862.

Trotter, H. F. [1957]. Gauss' work (1803-1826) on the theory of least squares. (English translation). Statistical Techniques Research Group, Department of Mathematics, Princeton University, Princeton.

Turner, M. E., Monroe, R. J., and Homer, L. D. [1963]. Generalized kinetic regression analysis: hypergeometric kinetics. Biometrics 19, 406-428.

Turner, M. E., Monroe, R. J., and Lucas, H. J. Jr. [1961]. Generalized asymptotic regression and non-linear path analysis. Biometrics 17, 120-143.

Wiggins, A. D. [1960]. On a multicompartment migration model with chronic feeding. Biometrics 16, 642-658

Wilks, S. S. [1962]. Mathematical Statistics. John Wiley and Sons, Incorporated, New York.

Zellner, A. [1962]. An efficient method of estimating seemingly unrelated regression and tests for aggregation bias. J. of the Amer. Statist. Assoc. 57, 348-368.